Trends in Cognitive Sciences

CellPress
REVIEWS

Review

Forms and Functions of the Self-Conscious Emotions

Daniel Sznycer [1,*]

Pride, shame, and guilt color our highest and lowest personal moments. Recent evidence suggests that these self-conscious emotions are neurocognitive adaptations crafted by natural selection. Specifically, self-conscious emotions solve adaptive problems of social valuation by promoting the achievement of valued actions and characteristics to increase others' valuations of the individual (pride); limiting information-triggered devaluation (shame); and remedying events where one put insufficient weight on the welfare of a valuable other (guilt). This adaptationist perspective predicts a form–function fit: a correspondence between the adaptive function of a self-conscious emotion and its information-processing structure. This framework can parsimoniously explain known facts about self-conscious emotions, make sense of puzzling findings, generate novel hypotheses, and explain why self-conscious emotions have their characteristic self-reflexive phenomenology.

What Are Self-Conscious Emotions?

Pride, shame, and guilt grace our successes and taint our failures. These **self-conscious emotions** (see Glossary) are not just feelings. These **emotions** motivate us to achieve, to avoid discredit, and to avoid harming those who are dear to us [1–5]. Indeed, self-conscious emotions are found beneath face-saving ploys, honor killings, wars, reconciliations, and achievements great and small [6–10].

Initially lagging relative to research on **basic emotions**, research on self-conscious emotions accelerated in the 1990s and early 2000s. This resurgence was spurred in great part by attributional theories; a paradigm that remains central to the study of these emotions. According to attributional theories, the activation and operation of pride, shame, and guilt depend critically on how the individual views and evaluates himself [11–14]. In this way, attributional theories highlight the intrapersonal nature of these emotions.

Nevertheless, basic questions about self-conscious emotions remain unanswered. Perhaps the central question is: why is the human mind/brain equipped with self-conscious emotions at all – what are these emotions for? Their dramatic interpersonal effects might suggest that self-conscious emotions have interpersonal adaptive functions. However, under attributional theories, interpersonal effects are secondary and even incidental to intrapersonal processes [15]. Consequently, the functions of self-conscious emotions remain puzzling.

Here, I review recent theory and data suggesting that self-conscious emotions have interpersonal adaptive functions and matching neurocognitive architectures realizing these functions. This interpersonal **adaptationist** framework can: (i) parsimoniously explain known facts about self-conscious emotions; (ii) make sense of puzzling findings; (iii) generate novel hypotheses,

Highlights

Self-conscious emotions such as pride, shame, and guilt are often studied through the lens of attributional theories. Under attributional theories, the activation and operation of self-conscious emotions depend on how the individual construes and evaluates her own successes and failures.

Although attributional theories highlight the intrapersonal nature of self-conscious emotions, recent theories and data suggest that the self-conscious emotions serve interpersonal adaptive functions.

From an adaptationist perspective, the characteristic self-reflexive and self-evaluative processes of self-conscious emotions are proximate means to solve adaptive problems related to social valuation.

Many known facts about the self-conscious emotions can be interpreted as outputs delivered by well-engineered emotion adaptations.

Attributional theories view shame as an immoral, pathological version of guilt. However, shame and guilt simply appear to be distinct adaptations serving different adaptive functions.

This interpersonal adaptationist framework can generate novel, testable hypotheses.

[1]Department of Psychology, University of Montreal, Montreal, QC H3C 3J7, Canada

*Correspondence: dsznycer2@gmail.com (D. Sznycer).

and (iv) explain why self-conscious emotions have their characteristic self-reflexive phenomenology.

Putting Self and Other into Self-Conscious Emotions

Pride, shame, and guilt are intrapersonal emotions. Self-reflexive and self-evaluative processes are key to understand these self-conscious emotions. This summarizes attributional theories, a paradigm that generated much of the existing research on these emotions [11–14]. Although emotions such as embarrassment, shyness, and social anxiety are also considered self-conscious emotions, the architectural nature of these emotions is not clear (see Outstanding Questions). For this reason, here, I focus on the more focal self-conscious emotions: pride, shame, and guilt.

According to attributional theories, self-conscious emotions trigger when events relevant to the individual's identity goals (e.g., the person one wants to be) are attributed to causes within the individual [11,12,14]. Additional **attributions** determine which specific emotion is elicited in a given situation. Guilt triggers when events deemed incongruent with one's identity goals (e.g., failing an exam, if one aspires to be a good student) are attributed to specific, unstable, or controllable aspects of the self (e.g., not having studied enough) [12,14]. Attributions to the global, stable, or uncontrollable self (e.g., being unintelligent) trigger shame instead [12,14]. Meanwhile, identity-goal-congruent outcomes (e.g., acing an exam) elicit achievement-oriented pride or hubris, depending on whether those outcomes are attributed to specific/unstable or global/stable aspects of the self [14,16]. Critically, under attributional theories, self-conscious emotions are intrapersonal emotions. The properties of these emotions, from their attribution-mediated elicitation to their affective and behavioral properties, sensitively depend on precisely how the individual construes and evaluates her own successes and failures.

However, recent theory and data suggest that self-conscious emotions have interpersonal adaptive functions [1,3,5,8,17–20] and matching neurocognitive architectures realizing these functions [1,21]. Self-conscious emotions appear to be information-processing **adaptations** tailored by **natural selection** because they helped our human ancestors navigate challenges and opportunities related to **social valuation** – the disposition to attend to others, associate with others, or trade personal welfare in favor of the welfare of others. These **adaptive problems** include: promoting and advertising achievements to increase others' valuations of the self (pride); limiting the spread of negative information about the self and any ensuing devaluation by others (shame); and remedying events where one put insufficient weight on the welfare of a valuable other (guilt). For alternative adaptationist theories of self-conscious emotions, see [3,17,22], Table 1, and Box 1. Over the millennia, individuals would have survived and reproduced to the degree that they navigated those challenges and opportunities effectively and efficiently. Natural selection would have differentially retained those neurocognitive variants (and their underlying genes) that solved those adaptive problems reliably. If so, the self-conscious emotions of contemporary humans should display a form–function fit; a close causal correspondence between the architecture of a self-conscious emotion, its adaptive function, and the statistical complex of ancestral regularities that caused that emotion to evolve [23]. Note that there are both agreements and disagreements between adaptationist and attributional theories of self-conscious emotions (Table 1).

The Evolution of Human Social Valuation

The evolutionary and cognitive roots of human social valuation are key to understand the self-conscious emotions. Humans evolved in a world of scarcity, disease, injury attacks by predators and conspecifics, and high mortality [24,25], and they relied on fellow group

Glossary

Adaptation: inherited part of an organism that became part of the standard design of the organism because it has reliably solved an adaptive problem throughout its evolution. While organisms also feature byproducts of adaptations and genetic and developmental noise, adaptations are the only parts that are adaptively functional.
Adaptationism: systematic analysis of adaptive design, or adaptations, in organisms.
Adaptive problem: evolutionarily recurrent task whose solution would have increased the likelihood of reproduction of an organism, however distally. For example: finding food, recognizing objects, and avoiding social devaluation.
Attribution: inference about the cause of an action or characteristic of self or others.
Basic emotion: emotion characterized by quick onset, brief duration, unbidden elicitation, and distinctive and universal elicitors and expression [119]. Basic emotions include: anger, fear, happiness, and disgust [119]. According to attributional theories, self-evaluation is standard in self-conscious emotions but merely optional in basic emotions [14].
Emotion: although definitions of this term vary [120], here, I define it as a particular type of adaptation that is designed to coordinate the operation of multiple different cognitive systems to solve complex adaptive problems [74]. For example: predator fear evolved to avoid predators; pride evolved to further the value of the self in the minds of others.
Internal regulatory variable: internal register of the value of a biologically relevant variable, which other cognitive mechanisms access for behavior regulation [40,121].
Natural selection: evolutionary process that retains those genetic variants that, in interaction with their relevant environments, reliably out-replicate alternative variants. Natural selection is the only evolutionary process that can produce complex organismic design–adaptations.
Recalibration: modification of an open parameter of the cognitive architecture in response to indications that a current setting

members for the assistance necessary to survive and reproduce. In this world, an individual would have thrived, struggled, or died early based on her ability to incentivize other group members to value her [26]; that is, to attend to the individual, to choose the individual as friend, mate, trading partner, and fellow coalition member, and to weight the individual's welfare when making decisions so that they would assist her when in need.

Different adaptive problems would have selected for cognitive mechanisms to value and help others. These adaptive problems include: helping one's kin; reciprocating goods and favors; managing one's reputation; pooling resources to smooth out variance in consumption; regulating one's exposure to the externalities emitted by fellow group members; choosing mates and social partners; and (substituting deference for valuation) claiming and defending resources by force [25,27–30]. These adaptive problems crafted specialized choice architectures to promote altruistic (or selfish) decisions given the information available to the actor about a potential recipient [31,32]. Both the ability to confer benefits (e.g., having skills) and the ability to aggressively inflict costs (e.g., being physically formidable) act as inputs to the systems that compute the social value of others and to the **internal regulatory variables** that dictate how much weight to attach to another's welfare based on their value to the individual.

Much of human sociality can be understood in terms of the operations of cognitive mechanisms that evolved to compute, store, **recalibrate**, and deploy the social valuations held by self and others. For example, the feeling of self-esteem appears to reflect an internal estimate of the degree to which others accept and include the self [33]. As expected, self-esteem closely tracks others' inclusion of the self [33,34]. Self-esteem and other internal indices of the individual's value to others (e.g., social status [35,36]) are coupled to emotional, motivational, and reasoning systems that function prospectively and reactively to optimize, within various constraints, others' valuations of the self. Jointly, these systems guide behavior to regain inclusion when one is excluded [37], to manage others' impressions of the self [38,39], and so forth.

In short, others' assessments of the acts and characteristics of a focal individual lead them to value (or disvalue) her. When others (an audience) detect new information about an individual that is at odds with their current level of valuation, their valuation is recalibrated either upward or downward, with correspondingly positive or negative effects on the individual's fitness [40]. Such shifts in social valuation constitute the proper domain of the self-conscious emotions.

Self-Conscious Emotions: Form Follows Function

Novel adaptive problems arise with the evolutionary appearance of conspecifics who are intrinsically valuable to the individual, and who can conditionally value or disvalue the individual based on the individual's actions and characteristics. For example, the value of a trading partner to an individual can be positive, null, or negative, depending on the former's ability and willingness to deliver valued goods to the latter. In contrast, a biological sibling is intrinsically valuable, because the replication prospects of your genes are enhanced when your sibling captures benefits or avoids costs.

Self-conscious emotions would have evolved as solutions to some adaptive problems of social valuation (for emotions solving other problems of social valuation [32,41–45]). Indeed, known facts about the self-conscious emotions can be interpreted in the light of particular adaptive problems of social valuation.

deviates from its actual or optimal value. For example, learning that your failure to help your friend occasioned her higher costs than you had anticipated can trigger a guilt-mediated revaluation of the cost imposed on her— – a recalibration that can lead you to help her subsequently. Recalibration is a key feature of self-conscious emotions [74].

Self-conscious emotion: emotion featuring self-reflexive mental processes that evolved to solve an adaptive problem of social valuation.

Social norm: normative standard of behavior that is enforced by a community (Box 1).

Social valuation: computational state that inclines the individual to attend to or associate with a specific other individual, or to trade her own welfare in favor of the welfare of that individual. Social valuation is subserved by functionally specialized internal regulatory variables and the motivational and behavioral mechanisms that access them [32,40].

Table 1. Similarities and Differences between Attributional and Adaptationist Theories of Self-Conscious Emotions

		Attributional theories	Adaptationist theories
		Similarities	
		Self-conscious emotions are emotion programs [120]; are elicited by (moral and nonmoral) successes and failures of the self; and elicit self-relevant cognition. Some self-conscious emotions (e.g., guilt) tend to have more socially desirable effects than others (e.g., shame).	
		Differences	
Focus		Intrapersonal	Interpersonal
Organizing representation		Ideal or current self-representation [13,14]	H.1: social value, social valuation [1,5,21] H.2: Social norm [3]
Source of organizing representation		Society [14], culture [13], and socialization [13].	Innate, domain-specific architecture of social valuation with invariant principles plus parameters open to local information [1,21,22,51,99].
Guilt[a]	Trigger	Attribution of event that is incongruent with one's identity goals to specific/controllable aspect of the self [12,14].	Insufficiently valuing a valuable other, independent of whether the other will know it [1,2,5,74].
	Adaptive function, effects[b]	Repairing relationship [107,14].	Increasing one's valuation of the other [1,2,5,74].
Shame[a]	Trigger	Attribution of event that is incongruent with one's identity goals to global/uncontrollable aspect of the self [12,14].	H.1: threat of being devalued due to spread of negative information about the self [1,5,65]. H.2: Violation of a social norm [3]. H.3: Interaction with dominant or higher-ranking other [3,17].
	Adaptive function, effects[b]	H.1: maladaptive, because of its association with aggression, paranoid thoughts, and depression [9,15,126]. Likely adaptive ancestrally, when dominance was a stronger determinant of status [13]. H.2: avoiding fitness costs of social rejection [14].	H.1: minimizing likelihood and costs of being devalued [1,5,65]. H.2: restoring conformity with violated norm [3]. H.3: avoiding subordinance [3] or attack by a formidable rival [17].
Pride[a]	Trigger	Attribution of event that is congruent with one's identity goals to specific/controllable or global/uncontrollable aspects of the self (triggering achievement-oriented pride or hubris, respectively) [12,14].	H.1: presence of opportunity to further the social value of the self in the minds of others [4,21]. H.2: fulfillment of a social norm [3]. H.3: interaction with submissive or lower-ranking other [3,17].
	Adaptive function, effects[b]	H.1: achievement-oriented pride; adaptive, because of prosocial effects [16]. H.2: hubris: maladaptive, because of antisocial effects [16]. (NB: elsewhere, achievement-oriented pride and hubris are viewed as adaptations for attaining prestige and dominance [96])	H.1: motivating and advertising the achievement of acts and characteristics that would increase others' valuation of the self [4,21]. H.2: rewarding conformity with social norms [3]. H.3: promoting dominance [3,17].

H, hypothesis. In some cases, more than one hypothesis has been advanced within a given paradigm; some but not all of these hypotheses are mutually exclusive.
[a]Some researchers view shame/guilt and pride as a single system [17].
[b]Attributional theories focus on the effects of self-conscious emotions; adaptationist theories focus on the adaptive functions of these emotions.

Pride

The emotion of pride appears to capitalize on opportunities to promote the social value of the individual in the minds of others. A system designed for this function should motivate the pursuit of acts or the cultivation of characteristics that others value (or fear). The system should also motivate the advertisement of those acts and characteristics, and exploit the enhanced social landscape that follows increases in the individual's ability to confer benefits or impose costs [3,8,17,21].

This theory of adaptive function can account for many known facts about pride. Pride-like behavior is taxonomically widespread [46], and therefore phylogenetically ancient. Pride is

Box 1. Social Norms and Self-Conscious Emotions

Some researchers study self-conscious emotions by reference to social norms. For example, it has been argued that violating a norm triggers shame, and fulfilling a norm triggers pride [3,122]. Once activated, shame and pride function to promote or reward conformity with social norms [3,22,123] in order to maintain access to the social benefits of cooperation and coordination.

Norm-based theories of self-conscious emotions tend to be observationally adequate. For example, the statement 'Scott feels shame because (people found that) he violated the norm against theft' makes intuitive sense. Furthermore, punishment can cause any type of behavior to be evolutionarily stable [124], and consistent with this, people moralize vast numbers of vastly different things. This makes norm-like general explanations appealing.

The lynchpin concept of norm is problematic, however. Common technical definitions of norm include, for example, 'cultural understandings concerning the normal, appropriate, or reasonable way to behave' [22]; 'normative standards of behavior that are enforced by informal social sanctions' [125]. These definitions are tautological, vague, or both. Indeed, existing social norms have little in common beyond their normativeness. For instance, because of kin selection, it is a cooperation norm to approach and help close kin; because of selection against inbreeding depression, it is a sex norm to avoid sex with close kin [32]. The concept of norm is superfluous when causal explanations are available, and it is a mere placeholder when explanations are not available yet. In either case, norm restates intuitions but fails to illuminate.

Definitions of norm do not happen to be vague; they are necessarily vague if vast numbers of different norms prescribing or proscribing different things in different domains are reduced to their common denominator.

Norm-based theories of self-conscious emotions face various problems. First, lumping all sources of shame or pride under the rubric norm obscures important differences. Consider the shame that arises from, for example, stinginess versus low productivity versus eating with the wrong fork. In theory, a well-designed shame system should discriminate functionally among antecedent conditions. In practice, it does [19,63].

Second, absent an ex ante, independently derived, and specific guide to know what is and what is not a norm, there is little to prevent one from deducing norms ex-post to explain observed occurrences of shame or pride. This invites circular reasoning and compromises the falsifiability of norm-based theories.

Third, hiding, lying, and worse are part of the modus operandi of shame [12,63,88,89,94,95]. Promoting conformity with norms cannot be the function of shame.

Fourth, norms are often thought to be culture specific. However, there are important crosscultural commonalities in what people value and disvalue in others, and in what elicits pride and shame [1,21,51,99]. Thus, these emotions may be governed less by culture-specific norms than by a species-wide architecture of social valuation comprised of invariant principles and open parameters.

triggered by aggressive formidability [47], achievements [8,48], and other socially valued characteristics. Pride is a highly pleasant emotion [49]; this internal reward can incentivize people to undertake and persevere at costly but socially valued courses of action [21,50,51]. Pride has a full-body display featuring an erect and expanded posture and gaze directed at the audience [3,48,52], and thus appears to generate common knowledge about the individual's enhanced value [53]. This display conveys achievement or dominance [3,17], is produced by congenitally blind individuals [47], and is recognized by young children [54] and by adults within and across cultures [55]. The pride display and related cues of being valued or feared have predictably functional effects on audiences. They appeal to potential mates [56], intimidate rivals [17], elicit submissiveness [57], and guide social learning through imitation [58].

Shame

Humans would have been selected to disvalue and shun individuals who are poor social partners [59,60]. This would have selected, on the recipient's end, for regulatory adaptations to minimize the spread of negative information about the self and the cost of any ensuing devaluation when negative information spreads [1,3,5,17].

Known facts about shame suggest that this emotion was engineered to counter devaluation. For example, when facing the prospect of being devalued, the individual inhibits actions that would cause others to devalue her [61,62]. The individual can also conceal or destroy incriminating information [63–65] and withdraw from the situation to avoid damage. Cues of being socially devalued elicit pain [66], which may deter devaluation-causing acts. When ashamed, the individual appeases [67] and produces a phylogenetically ancient [3,5,17] stereotyped nonverbal display [3,17,47] that deters attacks by signaling subordination; that is, that less weight on one's welfare is acceptable [68]. Compared with other displays (e.g., the anger display) or the absence of a display, the shame display mollifies observers of transgressions [67]. Social-evaluative threat upregulates proinflammatory cytokines [69] – advantageous when, for example, being physically punished results in infection. Experimental manipulations of prospective or actual devaluation reliably elicit shame [70–72], even for acts known by the individual to be irreproachable but mistakenly seen by others as violating a **social norm** [72] (Box 1).

While shame and pride aim to prevent or promote recalibrations of the valuations that others hold with respect to the individual, guilt recalibrates the individual's own valuations of others.

Guilt

When the reproductive fortune of an individual depends on that of another (as is the case among, e.g., kin, or friends), decision-making systems evolve to intrinsically value the other's welfare – not because there are benefits to be gained by conditionally cooperating, or costs to be avoided by propitiating the formidable, but because, within limits, enhancements of the other's welfare automatically (if indirectly) enhance the individual's own reproductive prospects [27,73]. Conversely, it is costly for an individual when a valuable other incurs costs or fails to obtain benefits. Therefore, it is a net cost for an individual when (i) she values the other's welfare less than what is dictated by the other's intrinsic value to her; or (ii) she underestimates how much the other values a good, service, or state of affairs, because then she will underdeliver those things.

Those costs can be abated through upward recalibrations of those variables. Although such revaluations would cause the individual to take more actions that benefit the other but cost her, the cost of the status quo is stipulated to be higher still. Thus, upward revaluations are the cost-effective alternative – up to the point where the incremental costs and incremental benefits of the revaluations equilibrate [40]. Importantly, when the other's welfare is intrinsically valuable to the individual, such revaluations should occur: (i) even when the other fails to protest or notice the individual's insufficient valuations; (ii) even when the other lacks the formidability to defend her interests (e.g., an infant); and (iii) even when there are no third parties that might devalue the individual. The guilt system appears to be the evolved solution to the adaptive problem of valuing insufficiently [2,70,74] (D. Sznycer, PhD thesis, University of California, Santa Barbara, 2010).

Consistent with this hypothesis, guilt tends to occur in the context of communal relationships [2,75]; that is, with respect to valuable interaction partners. Guilt interrupts the imposition of costs [75–77] and reduces re-offense [78]. Guilt motivates actions to benefit victims and repair relationships [2,75], including: restitutions, amends, apologies, confessions, perspective taking, and acceptance of responsibility [9,10,75,79–81]. Guilt is more limited in scope than shame is. Guilt is elicited in response to so-called moral failures (e.g., failing to help), but not in response to nonmoral failures (e.g., unattractiveness) [70,82]. Furthermore, guilt, unlike shame, is robustly elicited even when no one other than the perpetrator knows about the wrongdoing

[70]. Guilt predicts trustworthy behavior [83] and discourages partnerships with people who are more productive than the self, and who would therefore benefit one more than one would benefit them [84]. The situations in which guilt fails to mobilize are instructive. The induction of guilt increases altruistic behavior among dispositionally selfish people but not among dispositionally generous people [80,81]. Also, guilt activates following accidental rather than intentional transgressions [85]; that is, when the expression of a low interpersonal valuation falls short of the other's value to the individual, but not when the low valuation simply reflects the other's low value. Although, in dyadic situations, guilt leads to benefitting the victim at the expense of the self, when a third party is co-present, guilt can benefit the victim at the expense of the third party but not the self [86].

Within this framework, and in contrast with attributional theories [9,15,87], guilt is not the healthier substitute of shame; nor is guilt adaptive and shame ugly and maladaptive. Instead, these emotions are different regulatory programs that have evolved because they reliably solved different adaptive problems throughout human evolution. Guilt and shame trigger when the relevant cues meet the input conditions of either, or both, emotions. One can distinguish guilt and shame while seeing why they are related. In guilt, the outcome to be avoided is imposing undue harm on valuable others, even when the perpetrator faces no retaliation or reputational harm. In shame, the goal is to avoid being devalued by others. An act may elicit guilt and shame, but the eliciting conditions, computations, and outputs of these two systems are distinct. For example, someone who felt guilt and shame about infidelity might refrain from it, whereas someone who felt shame but not guilt about infidelity might practice it but conceal it.

Making Sense of Puzzling Facts

An interpersonal adaptationist approach can make sense of puzzling facts about the self-conscious emotions. Here I consider two of them.

The first puzzle concerns two inter-related aspects of shame: its adaptiveness and its effects. Because shame is associated with undesirable outcomes such as aggression [88,89], attributional researchers view shame as a maladaptive emotion [9,15]. However, this is perplexing, because maladaptive traits are edited out by the action of natural selection, and yet the shame system persists in the human mind/brain. Further complicating matters, shame motivates both antisocial [88,89] and prosocial [61,78,90] behaviors. For example, shame can motivate confessions and denials, approach and avoidance [90,91], and appeasement and externalization of blame [12,92,93]. This raises the question: why does shame deliver functionally antithetical behaviors?

These puzzles dissolve when considering that devaluation can be countered sometimes through prosocial means and sometimes through cunning and force. When prosocial behaviors are cost-effective means to rehabilitate one's social value in the eyes of others, shame will motivate them – a predicted response for a system designed to restore one's reputation as a good cooperative partner [65]. Otherwise, the shame system can switch to less noble means [88,89], which may be expected if social benefits are no longer as abundantly provided because of being valued but must instead be secured by deceit and aggression [3,17,31,60].

Recent findings support this hypothesis. For instance, the induction of shame causes dispositionally selfish people to cooperate more as second players in a sequential Prisoner's Dilemma game, where their defection would be uniquely traced to their selfishness. In contrast, in a simultaneous Prisoner's Dilemma game, where defection can also be attributed to a more benign fear of being defected on, shame does not cause more cooperation among the

dispositionally selfish [94]. That is, shame can inhibit defection, but shame allows defection when the situation affords cover.

Another example: when allocators in a money-allocation game offer little money to recipients who are ashamed, those recipients express less anger (compared with recipients in a no-shame control condition). This is so when the recipients know that the low allocators know why the recipients are ashamed. However, when recipients know that the low allocators do not know about the shame-causing event, the ashamed recipients express more anger [95]. In other words, the shame system tolerates poor treatment when others know about one's low social value, but poor treatment is angrily protested when it can be colorably portrayed as undeserved [41], because others do not know. In line with this conditional logic, a recent meta-analysis concluded: 'shame had a *positive* link to constructive approach when failure [. . .] or social image [. . .] was more reparable. In contrast, shame had a *negative* link to constructive approach when failure was less reparable' ([63], emphasis in original). Although a comprehensive decision tree of shame's behavior orchestration is yet to be elucidated, an adaptationist analysis of the existing evidence suggests that the outputs of shame are cost-effective and context-dependent means to defeat devaluation.

The fitness costs of being devalued are certainly 'ugly'. However, the shame system simply transmits devaluation-relevant information to other brain systems in order to minimize those costs.

The second puzzle concerns the affinity of pride with dominance. Indeed, dominance and aggression are hallmarks of hubristic pride [8]. However, this is odd, because pride is one of the most positively valenced emotions [49]. Why would such negative tactics as dominance and aggression taint, and even elicit [36], the status gains that fuel such a powerfully positive emotion as pride? As pride researchers have noted, 'Given the notably negative personality correlates of hubristic pride, it is not immediately evident why this facet would have evolved' [96]. Indeed, it has been argued that its link to aggression makes hubristic pride (also termed alpha pride) a maladaptive emotion [15,16].

Phylogenetic [3] and adaptationist [31] reasoning can shed light on this riddle. Cooperation has played a much greater role in the evolution of humans than in the evolution of non-human animals, and so others' positive fitness contributions to a valuer are correspondingly greater determinants of social valuation in humans. Nevertheless, humans retain and exploit phylogenetically ancient adaptations for aggression and dominance (including systems for assessment, threat, and fighting [3,17,31]). These adaptations can be deployed to obtain respect and status through intimidation, even when no positive contributions are made to fellow group members.

Consistent with the twin logic of human valuation via benefit delivery and cost imposition, both dominance and prosocial prestige predicted high status in a study of social interactions in groups. Participants did not like fellow group members who displayed dominance. Nonetheless, dominance, like prestige, was a viable route to influence and status [36]. Note that the reverse causal path from status and pride to dominance may operate as well, as gains in status may increase the cost-effectiveness of dominance as a means to obtain personal benefits [21,31,35].

Although dominance is undesirable to those on its receiving end and, often, to third parties, from the individual's perspective dominance can be a cost-effective tactic to incentivize deference from others. Hence its occurrence in pride.

Mapping Self-Conscious Emotions

Adaptationism can not only explain known facts about self-conscious emotions; this framework can also be used to generate novel, testable hypotheses.

If the structure of a self-conscious emotion echoes the statistical complex of regularities that over evolutionary time crafted that emotion, one can use knowledge and inferences about those ancestral regularities to uncover design features of self-conscious emotions. An application of this logic is described next.

Consider shame. A well-engineered shame system should mobilize not only reactively but also prospectively, before any devaluation occurs, in order to forestall actions that would cause others to devalue the individual [1,18,62,97]. To perform this prospective function, the anticipated magnitude of audience devaluation caused by a potential act (e.g., stealing) must be aversively weighted against the direct payoff of the act (e.g., acquiring). This would allow the individual to forego net costly acts in favor of net profitable acts. Importantly, this weighting and decision-making must often be made beforehand – the system would be severely handicapped if it needed to observe audience devaluation to compute its magnitude instead of estimating this magnitude in advance.

The devaluative audience psychology of contemporary humans is an approximate blueprint of the adaptive problem that shaped shame, and so it can be used to map the shame system and its anticipatory mode of operation.

t has been hypothesized that the anticipatory feeling of shame is an internally generated prediction that signals the magnitude of audience devaluation one would incur if one took an action that others devalue (scaled by the probability of detection) [1,98]. An internal shame signal precisely calibrated to forecast audience devaluation allows the individual to avoid two types of costly errors when devaluation looms: (i) shame underactivation, which leads to insufficient devaluation-minimizing measures and, therefore, excessive devaluation from others; and (ii) shame overactivation, which deters acts that yield more direct benefits than devaluation. This analysis implies the existence of a design feature: The shame system may be designed to forecast the precise magnitude of devaluation people in one's social ecology would experience if one took a given act that they disfavor, and deliver an internal shame signal whose intensity is proportional to it.

Experiments conducted in three industrial societies (USA, India, and Israel) supported this prediction. The intensity of anticipatory shame in every country closely tracked the magnitude of devaluation expressed by local audiences – in the absence of any communication between participants reporting their shame versus audiences reporting their devaluation in response to each of various acts [1] (e.g., stinginess, unattractiveness, poor table manners; Figure 1). Moreover, shame in every country also tracked the devaluation expressed by foreign audiences in the other two countries, suggesting universality not only in the structure of the shame system but also in its content [1]. Follow-up experiments indicated that audience devaluation is tracked specifically by shame, and not by other negatively valenced emotions [1]. Recently, it was found that shame tracks audience devaluation in each of 15 traditional small-scale societies [99], suggesting that shame's tracking of devaluation is a is a pan-human adaptation designed by selection, and not a product of cultural contact or convergent cultural evolution. Parallel though distinct predictions can be made regarding anticipatory guilt [100], although data on this point are missing.

Analogous reasoning suggests that anticipatory pride should forecast and track the degree to which audiences positively value different acts [101] in order to optimize the effectiveness–

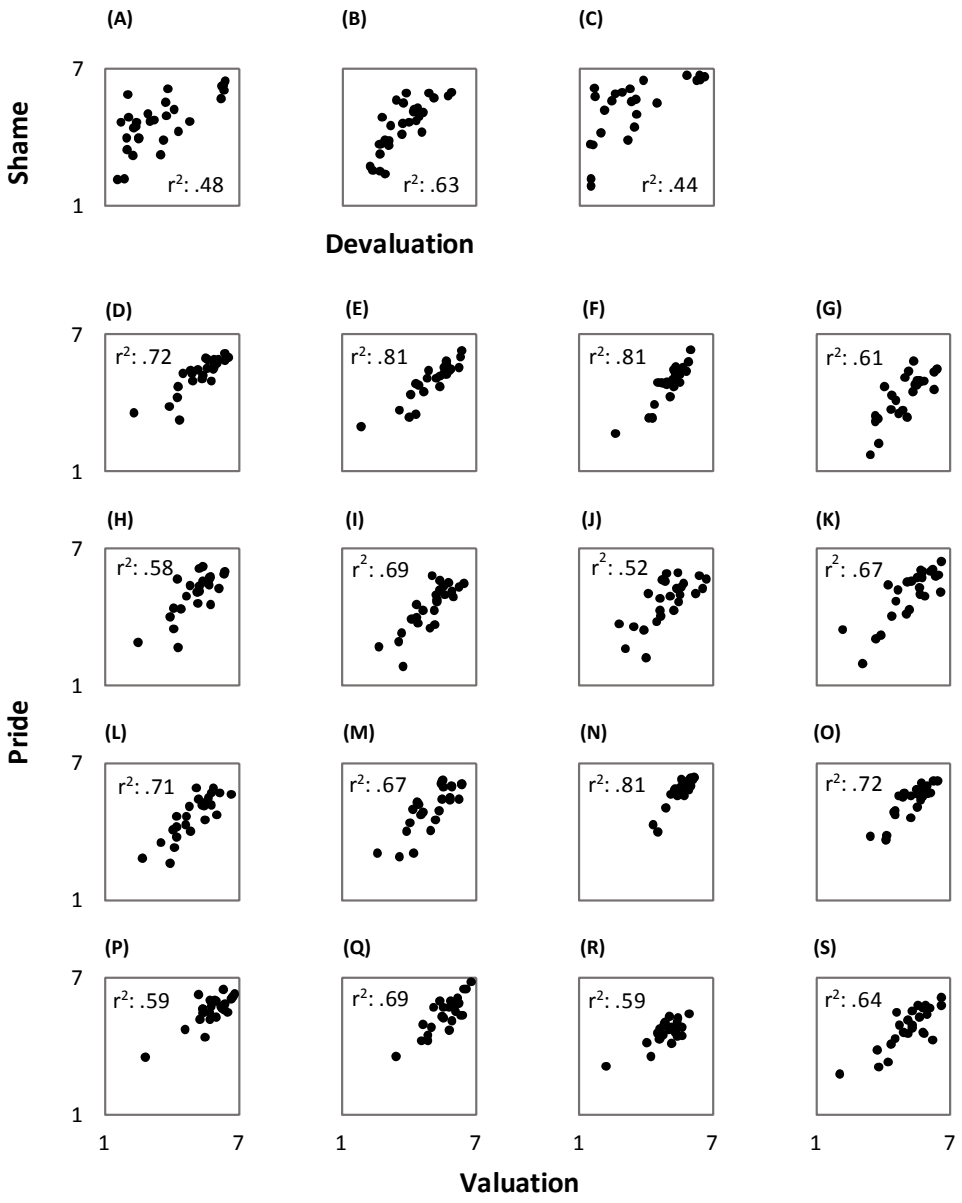

Figure 1. Shame and Pride Are Engineered to Balance the Competing Demands of Effectiveness and Efficiency. (A–C) Shame is designed to activate in proportion to the magnitude of audience devaluation (adapted from [1]); this minimizes the possible twin costs of underactivation and overactivation. Stimuli: hypothetical scenarios featuring discrediting situations, phrased from the perspective of the individual (e.g., 'You are not generous with others'; shame condition) or an observer (e.g., 'He is not generous with others'; audience condition) (between-subjects design). For each scenario, participants rated their feelings of shame if that situation was true of them (shame), and the negativity with which they would view the other person if that situation was true of the other – a measure of devaluation (audience). Each point represents the mean shame rating and mean devaluation rating of one scenario. Data from (N of scenarios): A: USA [29], B: India [29], C: Israel [24]. D–S: Pride tracks audience valuation (adapted from [21]). These studies were similar to the shame-devaluation studies, but here participants rated scenarios designed to elicit positive valuation from others. For each scenario, participants rated their pride (e.g., 'You finished first in a marathon'; pride condition) or their valuation of another individual (e.g., 'He finished first in a marathon'; audience condition). Each point represents the mean pride rating and

(Figure legend continued on the bottom of the next page.)

efficiency tradeoff in the pursuit of valuation from others. Data from industrial [21] and small-scale societies [51] support this prediction (Figure 1).

In summary, it is ineffective when shame and pride underactivate, and it is inefficient when these emotions overactivate. To avoid these pitfalls, shame and pride obey Goldilocks' 'just right' principle. Shame and pride can balance the competing demands of effectiveness and efficiency because they are adaptations engineered to activate in close proportion to the evaluations of audiences.

Interpersonal Aspects of Self-Consciousness

Self-conscious mentation is central to attributional theories' accounts of self-conscious emotions. As attributional researchers have remarked, these emotions 'are evoked by self-reflection and self-evaluation' [15]. Meanwhile, adaptationist thinking suggests that self-conscious emotions have interpersonal adaptive functions. Are these views incompatible?

Much of self-consciousness is not social [102]. However, might the self-consciousness of self-conscious emotions be means to interpersonal ends? Self-consciousness must somehow have generated the fitness benefits that would have fueled its continued replication over evolutionary time, and interpersonal functions constitute possible causal paths.

Consider shame. It is plausible that the self-consciousness of shame is projected by various recalibrational and decision-making procedures designed to counter devaluation. The following are a few possible candidates.

Based on indications about which member of the audience might know what, an early decision faced by the shame system is whether to mount a comprehensive response or feign normalcy – displaying shame can sometimes be a telltale sign of culpability. In either case, the shame system may upgrade precautions to escape detection in the future. Additionally, when mounting a response, the shame system may upregulate the weight attached to the welfare of others (at least in public); downgrade estimates of the value of one's own welfare to others [33]; and correspondingly downgrade the level of entitlement one should display thenceforth. These internal recalibrations may have a self-conscious phenomenology.

The characteristic self-blaming of shame too may support interpersonal functions. For instance, it has been argued that self-blame can be a self-protective tactic: a signal of submission to deter attacks [103]. Consistent with this, women in abusive relationships blame themselves, but they blame their partners when they exit the relationship [103]. Self-blame may also function to probe for others' assent (or dissent), and thus to gauge changes in others' evaluations of the self; to elicit sympathy and forgiveness; and to feign lack of ability to exempt oneself from future responsibilities [104–106].

An interpersonal approach can also explain differences between the self-conscious experiences afforded by shame and guilt – an object of much attributional research [12–14,107]. Recall that, under attributional theories, negative outcomes are attributed to the global/stable/uncontrollable self in shame, and to specific/unstable/controllable aspects of the self in guilt. In contrast, an adaptationist perspective suggests that those differences do not merely reflect

mean valuation rating of one scenario. Data from: D, USA; E, Canada; F, UK; G, France; H, Belgium; I, The Netherlands; J, Switzerland; K, Italy; L, Turkey; M, Israel; N, India; O, Singapore; P, Philippines; Q, South Korea; R, Japan; S, Australia. Number of scenarios = 25.

how the individual construes self-relevant events. Rather, those differences reflect real architectural differences between the shame and guilt programs, which in turn echo the contrasting adaptive problems posed by devaluation versus the insufficient valuation of valuable others [22]. Countering another's devaluation (shame) is less within the individual's control than upregulating one's valuation of another's welfare (guilt). This may explain the feeling that shame is less controllable than guilt is. Second, as stated above, guilt follows unintentional expressions of low valuation [85], whereas shame follows both unintentional and intentional outcomes (e.g., being physically unattractive vs stealing [1]). This may explain the intuition that shame events are more stably diagnostic of the individual's constitution than guilt events are. Third, shame-triggering events tend to have broader interpersonal ramifications than guilt-triggering events do. Consider failing to help a friend in need, an omission that may trigger both shame and guilt. The adaptive problem handled by guilt is automatically solved, and the operation of guilt is interrupted, once the individual effects the requisite recalibrations. In contrast, the failure to help may cause devaluation among the friend and among third parties indirectly affected by one's omission [108]. Thus, the problem handled by shame is not necessarily solved once the directly affected party undoes her devaluation. This may explain the feeling that shame is more global than guilt is. Finally, people attach more weight to negative (versus positive) information about others [109,110], so negative information about the self in the minds of others can feel global and stable.

Concluding Remarks and Future Directions

What is complex and functional in nervous systems has evolved because of an evolutionary history of contributing to the regulation of physiology and behavior in reproduction-promoting ways. For that reason, adaptationist thinking and knowledge of ancestral humans' physical and social ecologies are invaluable to reverse-engineering the brain and its computations. Current adaptationist theories of self-conscious emotions are necessarily tentative (see Outstanding Questions). Even so, these theories can parsimoniously explain many known facts about these emotions and guide the discovery of new features.

Adaptationist theories of self-conscious emotions are highly generative. The following are some hypotheses that might be profitably assessed in future work. First, countering devaluation is an inherently complex challenge, and the problem space of shame is correspondingly large (D. Sznycer, PhD thesis, University of California, Santa Barbara, 2010). Consider: feigning normalcy is a best response when others have not caught you red-handed yet, but not when they have. Confessing is a best response when information about your ignoble actions may have leaked, but not when everyone is in the dark [64]. Tolerating a reduction in status is a best response when most find your actions disgraceful, but not when opinions are divided or when you enjoy the backing of high-status allies. Functional thinking is expected to shed light on the behavioral repertoire of shame [19,63,111] (D. Sznycer, PhD thesis, University of California, Santa Barbara, 2010). Second, the fact that pride and shame track others' social valuations suggests that other emotions, self-conscious and non-self-conscious, may do so as well. The fact that pride and shame track others' valuations also suggests that those emotions interface informationally, and physically, with (at least) the individual's own valuation architecture. Consistent with this, some brain areas involved in the computation of social valuation (e.g., ventromedial prefrontal cortex/orbitofrontal cortex) [112] appear to also support the operation of self-conscious emotions [113–115]. Third, it has been argued that human institutions are underwritten by evolved intuitions that demarcate the social arrangements seen as normal from those seen as peculiar or senseless [116–118]. Self-conscious emotions and the valuation architecture appear to play an important role in the creation of human institutions and culture (e.g., the criminal justice system, the advertising industry).

Outstanding Questions

Why do self-conscious emotions vary across situations, over the lifespan, and between individuals and populations? Part of this variation may be functional. These emotions may feature invariant principles plus open parameters filled with local information – an architecture that confers functionality and flexibility. For example, shame, hubristic pride, and embarrassment peak during adolescence, a time of intense interactions with peers. This and other causes of variation in self-conscious emotions merit further investigation.

How many functionally distinct neural programs do the terms shame, guilt, embarrassment, shyness, and social anxiety refer to? Naturally occurring emotion terms are too blunt to carve up self-conscious emotions at their functional joints. More direct evidence from behavior, psychopathology, and neurophysiology suggests real differences between, for example, shame and guilt, and between shame and embarrassment, although the latter difference is subtler. Some of these terms may denote different functional architectures; others may denote different parameterizations of the same architecture; others may have no distinct counterpart.

How is the folk concept of AUDIENCE structured? While audience effects on cooperation, punishment, and other behaviors have been studied extensively, little is known about people's implicit concept of AUDIENCE. The psychology that generates concepts would have been selected to assist the individual regulate her behavior. It is speculated that the AUDIENCE concept evolved to strategically manage the flow of personal information to fellow group members. Simple physical or mental state descriptions may be insufficient to characterize AUDIENCE; for example, individuals of a given social status or knowledge state may not be tagged as AUDIENCE, even when copresent. Knowledge about self-conscious emotions will be helpful in dissecting this concept.

There are various agreements between adaptationist theories and attributional theories. Importantly, some disagreements are apparent and simply reflect differences in the focus of analysis. For example, the contrasting self-reflexive phenomenology of shame and guilt elucidated by intrapersonally oriented attributional theories is consistent with the contrasting interpersonal functions of those emotions hypothesized by adaptationist theories (see above). However, actual disagreements remain, between interpersonal/adaptationist and intrapersonal/attributional theories [15,16,18,22,70,101], and between different adaptationist theories [1,3].

However those disagreements are resolved, it is clear that self-conscious emotions are powerful motivators of human behavior. We now know a lot about what self-conscious emotions do. Next-generation models can profitably focus on how and why self-conscious emotions do what they do. Comprehensively mapping the information-processing structure of self-conscious emotions is a central task. Once that is done, rapid progress should follow. Light will be shed on how self-conscious emotions develop, how they are instantiated algorithmically and physically, how they fail in clinical populations, and how they give rise to personality, sex, and cultural differences.

Acknowledgments

I thank Laith Al-Shawaf, Lindsey Drayton, Tadeg Quillien, and two anonymous reviewers for helpful comments on an earlier version of this manuscript.

References

1. Sznycer, D. *et al.* (2016) Shame closely tracks the threat of devaluation by others, even across cultures. *Proc. Natl. Acad. Sci. U. S. A.* 113, 2625–2630
2. Baumeister, R.F. *et al.* (1994) Guilt: An interpersonal approach. *Psychol. Bull.* 115, 243–267
3. Fessler, D.M.T. (1999) Toward an understanding of the universality of second order emotions. In *Beyond Nature or Nurture: Biocultural Approaches to the Emotions* (Hinton, A.L., ed.), pp. 75–116, Cambridge University Press
4. Tracy, J.L. *et al.* (2010) A naturalist's view of pride. *Emot. Rev.* 2, 163–177
5. Gilbert, P. (1997) The evolution of social attractiveness and its role in shame, humiliation, guilt and therapy. *Br. J. Med. Psychol.* 70, 113–147
6. Nisbett, R.E. and Cohen, D. (1996) *Culture of Honor: The Psychology of Violence in the South,* Westview
7. Hobbes, T. (1651/1985) *Leviathan*, Penguin
8. Tracy, J.L. (2016) *Take Pride: Why the Deadliest Sin Holds the Secret to Human Success,* Houghton Mifflin Harcourt
9. Tangney, J.P. (1991) Moral affect: the good, the bad, and the ugly. *J. Pers. Soc. Psychol.* 61, 598–607
10. Leith, K.P. and Baumeister, R.F. (1998) Empathy, shame, guilt, and narratives of interpersonal conflicts: Guilt-prone people are better at perspective taking. *J. Pers.* 66, 1–37
11. Lewis, H.B. (1971) *Shame and Guilt in Neurosis,* International Universities Press
12. Tangney, J.P. (1990) Assessing individual differences in proneness to shame and guilt: development of the Self-Conscious Affect and Attribution Inventory. *J. Pers. Soc. Psychol.* 59, 102–111
13. Tangney, J.P. and Dearing, R.L. (2002) *Shame and Guilt,* Guilford Press
14. Tracy, J.L. and Robins, R.W. (2004) Putting the self into self-conscious emotions: a theoretical model. *Psychol. Inq.* 15, 103–125
15. Tangney, J.P. *et al.* (2007) Moral emotions and moral behavior. *Annu. Rev. Psychol.* 58, 345–372
16. Tracy, J.L. and Robins, R.W. (2014) Conceptual and empirical strengths of the authentic/hubristic model of pride. *Emotion* 14, 33–37
17. Weisfeld, G.E. and Dillon, L.M. (2012) Applying the dominance hierarchy model to pride and shame, and related behaviors. *J. Evol. Psychol.* 10, 15–41
18. Leary, M.R. (2015) Emotional responses to interpersonal rejection. *Dialogues Clin. Neurosci.* 17, 435–441
19. de Hooge, I.E. *et al.* (2011) A functionalist account of shame-induced behaviour. *Cogn. Emot.* 25, 939–946
20. Rodriguez Mosquera, P.M. (2018) Cultural concerns: How valuing social-image shapes social emotion. *Eur. Rev. Soc. Psychol.* http://dx.doi.org/10.1080/10463283.2017.1412180
21. Sznycer, D. *et al.* (2017) Cross-cultural regularities in the cognitive architecture of pride. *Proc. Natl. Acad. Sci. U. S. A.* 114, 1874–1879
22. Fessler, D.M.T. *et al.* (2007) From appeasement to conformity: Evolutionary and cultural perspectives on shame, competition, and cooperation. In *The Self-Conscious Emotions: Theory and Research* (Tracy, J.L., ed.), pp. 174–193, Guilford Press
23. Barrett, H.C. (2014) *The Shape of Thought: How Mental Adaptations Evolve,* Oxford University Press
24. Burger, O. *et al.* (2012) Human mortality improvement in evolutionary context. *Proc. Natl. Acad. Sci. U. S. A.* 109, 18210–18214
25. Hill, K. and Hurtado, A.M. (1996) *Ache Life History: the Ecology and Demography of a Foraging People,* de Gruyter
26. von Rueden, C. and Jaeggi, A.V. (2016) Men's status and reproductive success in 33 nonindustrial societies: effects of subsistence, marriage system, and reproductive strategy. *Proc. Natl. Acad. Sci. U. S. A.* 113, 10824–10829
27. Tooby, J. and Cosmides, L. *et al.* (1996) Friendship and the Banker's Paradox: other pathways to the evolution of adaptations for altruism. In *Evolution of Social Behaviour Patterns in Primates and Man. Proceedings of the British Academy 88* (Runciman, W.G., ed.), pp. 119–143
28. Hammerstein, P. and Parker, G.A. (1982) The asymmetric war of attrition. *J. Theor. Biol.* 96, 647–682
29. Nowak, M.A. (2006) Five rules for the evolution of cooperation. *Science* 314, 1560–1563
30. Sober, E. and Wilson, D.S. (1999) *Unto Others: The Evolution and Psychology of Unselfish Behavior,* Harvard University Press

31. Sell, A. *et al.* (2016) Bargaining power and adolescent aggression: the role of fighting ability, coalitional strength, and mate value. *Evol. Hum. Behav.* 37, 105–116
32. Lieberman, D. *et al.* (2007) The architecture of human kin detection. *Nature* 445, 727–731
33. Leary, M.R. *et al.* (1995) Self-esteem as an interpersonal monitor: the sociometer hypothesis. *J. Pers. Soc. Psychol.* 68, 518–530
34. Eisenberger, N.I. *et al.* (2011) The neural sociometer: brain mechanisms underlying state self-esteem. *J. Cogn. Neurosci.* 23, 3448–3455
35. Mahadevan, N. *et al.* (2018) Is self-regard a sociometer or a hierometer? Self-esteem tracks status and inclusion, narcissism tracks status. *J. Pers. Soc. Psychol.* http://dx.doi.org/10.1037/pspp0000189
36. Cheng, J.T. *et al.* (2013) Two ways to the top: evidence that dominance and prestige are distinct yet viable avenues to social rank and influence. *J. Pers. Soc. Psychol.* 104, 103–125
37. Williams, K.D. and Nida, S.A. (2011) Ostracism: consequences and coping. *Curr. Dir. Psychol. Sci.* 20, 71–75
38. Sedikides, C. and Strube, M.J. (1997) Self-evaluation: to thine own self be good, to thine own self be sure, to thine own self be true, and to thine own self be better. *Adv. Exp. Soc. Psychol.* 29, 209–269
39. Leary, M.R. and Kowalski, R.M. (1990) Impression management: a literature review and two-component model. *Psychol. Bull.* 107, 34–47
40. Tooby, J. *et al.* (2008) Internal regulatory variables and the design of human motivation: a computational and evolutionary approach. In *Handbook of Approach and Avoidance Motivation* (Elliot, A.J., ed.), pp. 251–271, Lawrence Erlbaum Associates
41. Sell, A. *et al.* (2017) The grammar of anger: mapping the computational architecture of a recalibrational emotion. *Cognition* 168, 110–128
42. Al-Shawaf, L. *et al.* (2016) Human emotions: an evolutionary psychological perspective. *Emot. Rev.* 8, 173–186
43. Smith, A. *et al.* (2017) Cooperation: the roles of interpersonal value and gratitude. *Evol. Hum. Behav.* 38, 695–703
44. Sznycer, D. *et al.* (2017) Support for redistribution is shaped by compassion, envy, and self-interest, but not a taste for fairness. *Proc. Natl. Acad. Sci. U. S. A.* 114, 8420–8425
45. Sznycer, D. *et al.* (2017) Adaptationism carves emotions at their functional joints. *Psychol. Inq.* 28, 56–62
46. Huntingford, F. and Turner, A. (1987) *Animal Conflict,* Chapman and Hall
47. Tracy, J.L. and Matsumoto, D. (2008) The spontaneous display of pride and shame: evidence for biologically innate nonverbal displays. *Proc. Natl. Acad. Sci. U. S. A.* 105, 11655–11660
48. Lewis, M. *et al.* (1992) Differences in shame and pride as a function of children's gender and task difficulty. *Child Dev.* 63, 630–638
49. Mauro, R. *et al.* (1992) The role of appraisal in human emotions: a cross-cultural study. *J. Pers. Soc. Psychol.* 62, 301–317
50. Williams, L.A. and DeSteno, D. (2008) Pride and perseverance: the motivational role of pride. *J. Pers. Soc. Psychol.* 94, 1007–1017
51. Sznycer, D. *et al.* (2018) Invariances in the architecture of pride across small-scale societies. *Proc. Natl. Acad. Sci. U. S. A.* 115, 8322–8327
52. Tracy, J.L. and Robins, R.W. (2007) The prototypical pride expression: development of a nonverbal behavior coding system. *Emotion* 7, 789–801
53. Thomas, K.A. *et al.* (2018) Common knowledge, coordination, and the logic of self-conscious emotions. *Evol. Hum. Behav.* 39, 179–190
54. Tracy, J.L. *et al.* (2005) Can children recognize pride? *Emotion* 5, 251–257
55. Tracy, J.L. and Robins, R.W. (2008) The nonverbal expression of pride: evidence for cross-cultural recognition. *J. Pers. Soc. Psychol.* 94, 516–530
56. de Bruyn, E.H. *et al.* (2012) Dominance-popularity status, behavior, and the emergence of sexual activity in young adolescents. *Evol. Psychol.* 10, 296–319
57. Tiedens, L.Z. and Fragale, A.R. (2003) Power moves: complementarity in dominant and submissive nonverbal behavior. *J. Pers. Soc. Psychol.* 84, 558
58. Chudek, M. *et al.* (2012) Prestige-biased cultural learning: bystander's differential attention to potential models influences children's learning. *Evol. Hum. Behav.* 33, 46–56
59. Kurzban, R. and Leary, M.R. (2001) Evolutionary origins of stigmatization: the functions of social exclusion. *Psychol. Bull.* 127, 187–208
60. Hales, A.H. *et al.* (2016) Disagreeableness as a cause and consequence of ostracism. *Pers. Soc. Psychol. Bull.* 42, 782–797
61. de Hooge, I.E. *et al.* (2008) Not so ugly after all: when shame acts as a commitment device. *J. Pers. Soc. Psychol.* 95, 933–943
62. Fehr, E. and Gächter, S. (2000) Cooperation and punishment in public goods experiments. *Am. Econ. Rev.* 90, 980–994
63. Leach, C.W. and Cidam, A. (2015) When is shame linked to constructive approach orientation? A meta-analysis. *J. Pers. Soc. Psychol.* 109, 983–1002
64. Sznycer, D. *et al.* (2015) Regulatory adaptations for delivering information: the case of confession. *Evol. Hum. Behav.* 36, 44–51
65. de Hooge, I.E. *et al.* (2010) Restore and protect motivations following shame. *Cogn. Emot.* 24, 111–127
66. Eisenberger, N.I. *et al.* (2003) Does rejection hurt? An fMRI study of social exclusion. *Science* 302, 290–292
67. Keltner, D. *et al.* (1997) Appeasement in human emotion, social practice, and personality. *Aggress. Behav.* 23, 359–374
68. Gilbert, P. (2000) The relationship of shame, social anxiety and depression: the role of the evaluation of social rank. *Clin. Psychol. Psychother.* 7, 174–189
69. Dickerson, S.S. *et al.* (2009) Social-evaluative threat and proinflammatory cytokine regulation. *Psychol. Sci.* 20, 1237–1244
70. Smith, R.H. *et al.* (2002) The role of public exposure in moral and nonmoral shame and guilt. *J. Pers. Soc. Psychol.* 83, 138–159
71. Dickerson, S.S. *et al.* (2008) Negative social evaluation, but not mere social presence, elicits cortisol responses to a laboratory stressor task. *Health Psychol.* 27, 116–121
72. Robertson, T.E. *et al.* (2018) The true trigger of shame: social devaluation is sufficient, wrongdoing is unnecessary. *Evol. Hum. Behav.* 39, 566–573
73. Tomasello, M. *et al.* (2012) Two key steps in the evolution of human cooperation: the interdependence hypothesis. *Curr. Anthropol.* 53, 673–692
74. Tooby, J. and Cosmides, L. *et al.* (2008) The evolutionary psychology of the emotions and their relationship to internal regulatory variables. In *Handbook of Emotions* (Lewis, M., ed.), pp. 114–137, Guilford
75. Baumeister, R.F. *et al.* (1995) Personal narratives about guilt: role in action control and interpersonal relationships. *Basic Appl. Soc. Psychol.* 17, 173–198
76. Cohen, T.R. *et al.* (2014) Moral character in the workplace. *J. Pers. Soc. Psychol.* 107, 943
77. Cohen, T.R. *et al.* (2013) Predicting counterproductive work behavior from guilt proneness. *J. Bus. Ethics* 114, 45–53
78. Tangney, J.P. *et al.* (2014) Two faces of shame: the roles of shame and guilt in predicting recidivism. *Psychol. Sci.* 25, 799–805
79. Ohtsubo, Y. and Yagi, A. (2015) Relationship value promotes costly apology-making: Testing the valuable relationships hypothesis from the perpetrator's perspective. *Evol. Hum. Behav.* 36, 232–239
80. de Hooge, I.E. *et al.* (2007) Moral sentiments and cooperation: differential influences of shame and guilt. *Cogn. Emot.* 21, 1025–1042

81. Ketelaar, T. and Au, W.T. (2003) The effects of feelings of guilt on the behaviour of uncooperative individuals in repeated social bargaining games: an affect-as-information interpretation of the role of emotion in social interaction. *Cogn. Emot.* 17, 429–453
82. Giner-Sorolla, R. and Espinosa, P. (2011) Social cuing of guilt by anger and of shame by disgust. *Psychol. Sci.* 22, 49–53
83. Levine, E.E. *et al.* (2018) Who is trustworthy? Predicting trustworthy intentions and behavior. *J. Pers. Soc. Psychol.* 115, 468–494
84. Wiltermuth, S.S. and Cohen, T.R. (2014) "I'd only let you down": guilt proneness and the avoidance of harmful interdependence. *J. Pers. Soc. Psychol.* 107, 925–942
85. McGraw, K.M. (1987) Guilt following transgression: an attribution of responsibility approach. *J. Pers. Soc. Psychol.* 53, 247–256
86. de Hooge, I.E. *et al.* (2011) What is moral about guilt? Acting "prosocially" at the disadvantage of others. *J. Pers. Soc. Psychol.* 100, 462–473
87. Leach, C.W. (2017) Understanding shame and guilt. In *Handbook of the Psychology of Self-Forgiveness*, pp. 17–28, Springer
88. Tangney, J.P. *et al.* (1992) Shamed into anger? The relation of shame and guilt to anger and self-reported aggression. *J. Pers. Soc. Psychol.* 62, 669–675
89. Elison, J. *et al.* (2014) Shame and aggression: theoretical considerations. *Aggress. Violent Behav.* 19, 447–453
90. Gausel, N. *et al.* (2016) Resolving the paradox of shame: differentiating among specific appraisal-feeling combinations explains pro-social and self-defensive motivation. *Motiv. Emot.* 40, 118–139
91. de Hooge, I.E. *et al.* (2018) The social side of shame: approach versus withdrawal. *Cogn. Emot.* http://dx.doi.org/10.1080/02699931.2017.1422696
92. Griffin, B.J. *et al.* (2016) Perpetrators' reactions to perceived interpersonal wrongdoing: The associations of guilt and shame with forgiving, punishing, and excusing oneself. *Self Identity* 15, 650–661
93. Gilbert, P. (2000) Varieties of submissive behavior as forms of social defense: their evolution and role in depression. In *Subordination and Defeat: An Evolutionary Approach to Mood Disorders and Their Therapy* (Sloman, L. and Gilbert, P., eds), Lawrence Erlbaum Associates
94. Declerck, C.H. *et al.* (2014) No place to hide: when shame causes proselfs to cooperate. *J. Soc. Psychol.* 154, 74–88
95. Zhu, R. *et al.* (2018) The effect of shame on anger at others: awareness of the emotion-causing events matters. *Cogn. Emot.* http://dx.doi.org/10.1080/02699931.2018.1489782
96. Cheng, J.T. *et al.* (2010) Pride, personality, and the evolutionary foundations of human social status. *Evol. Hum. Behav.* 31, 334–347
97. Somerville, L.H. *et al.* (2013) The medial prefrontal cortex and the emergence of self-conscious emotion in adolescence. *Psychol. Sci.* 24, 1554–1562
98. Crockett, M.J. *et al.* (2017) Moral transgressions corrupt neural representations of value. *Nat. Neurosci.* 20, 879
99. Sznycer, D. *et al.* (2018) Cross-cultural invariances in the architecture of shame. *Proc. Natl. Acad. Sci. U. S. A.* 115, 9702–9707
100. Frank, R.H. (1988) *Passions within Reason: The Strategic Role of the Emotions,* Norton
101. Holbrook, C. *et al.* (2014) Conceptual and empirical challenges to the "authentic" versus "hubristic" model of pride. *Emotion* 14, 17–32
102. Guterstam, A. *et al.* (2015) Posterior cingulate cortex integrates the senses of self-location and body ownership. *Curr. Biol.* 25, 1416–1425
103. Gilbert, P. (2016) *Depression: The Evolution of Powerlessness,* Routledge
104. Driscoll, R. (1989) Self-condemnation: A comprehensive framework for assessment and treatment. *Psychot. Theor. Res. Pract.* 26, 104
105. Tanaka, H. *et al.* (2015) Shame-prone people are more likely to punish themselves: a test of the reputation-maintenance explanation for self-punishment. *Evol. Behav. Sci.* 9, 1
106. Zhu, R. *et al.* (2017) Self-punishment promotes forgiveness in the direct and indirect reciprocity contexts. *Psychol. Rep.* 120, 408–422
107. Tangney, J.P. *et al.* (1996) Are shame, guilt, and embarrassment distinct emotions? *J. Pers. Soc. Psychol.* 70, 1256–1269
108. Pedersen, E.J. *et al.* (2018) The unresponsive avenger: more evidence that disinterested third parties do not punish altruistically. *J. Exp. Psychol. Gen.* 147, 514–544
109. Baumeister, R.F. *et al.* (2001) Bad is stronger than good. *Rev. Gen. Psychol.* 5, 323–370
110. Siegel, J.Z. *et al.* (2018) Beliefs about bad people are volatile. *Nat. Hum. Behav.* 2, 750
111. Lukaszewski, A.W. Evolutionary perspectives on the mechanistic underpinnings of personality. In *The Handbook of Personality Dynamics and Processes* (Rauthmann, J. F., ed), Elsevier (in press)
112. Levy, D.J. and Glimcher, P.W. (2012) The root of all value: a neural common currency for choice. *Curr. Opin. Neurobiol.* 22, 1027–1038
113. Beer, J.S. *et al.* (2006) Orbitofrontal cortex and social behavior: integrating self-monitoring and emotion-cognition interactions. *J. Cogn. Neurosci.* 18, 871–880
114. Darby, R.R. *et al.* (2018) Lesion network localization of criminal behavior. *Proc. Natl. Acad. Sci. U. S. A.* 115, 601–606
115. Bastin, C. *et al.* (2016) Feelings of shame, embarrassment and guilt and their neural correlates: A systematic review. *Neurosci. Biobehav. Rev.* 71, 455–471
116. Cosmides, L. and Tooby, J. (2006) Evolutionary psychology, moral heuristics, and the law. In *Heuristics and the Law* (Gigerenzer, G. and Engel, C., eds), pp. 175–205, MIT Press
117. Lieberman, D. and Patrick, C. (2018) *Objection: Disgust, Morality, and the Law,* Oxford University Press
118. Boyer, P. and Petersen, M.B. (2012) The naturalness of (many) social institutions: evolved cognition as their foundation. *J. Instit. Econ.* 8, 1–25
119. Ekman, P. (1992) An argument for basic emotions. *Cog. Emot.* 6, 169–200
120. Barrett, L.F. and Russell, J.A. (2014) *The Psychological Construction of Emotion,* Guilford Publications
121. Balliet, D. *et al.* (2017) Functional interdependence theory: an evolutionary account of social situations. *Pers. Soc. Psychol. Rev.* 21, 361–388
122. Elster, J. (1989) *The Cement of Society: a Survey of Social Order,* Cambridge University Press
123. Bicchieri, C. (2005) *The Grammar of Society: the Nature and Dynamics of Social Norms,* Cambridge University Press
124. Boyd, R. and Richerson, P.J. (1992) Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethol. Sociobiol.* 13, 171–195
125. Fehr, E. and Fischbacher, U. (2004) Third-party punishment and social norms. *Evol. Hum. Behav.* 25, 63–87
126. Tangney, J.P. *et al.* (1992) Proneness to shame, proneness to guilt, and psychopathology. *J. Abnorm. Psychol.* 101, 469–478