

Elster's Eclecticism in Analyzing Emotion¹

Inquiry, 64(3), 321-341 (2021). doi.org/10.1080/0020174x.2020.1821957

G. Ainslie
Department of Veterans Affairs, Coatesville, PA 19320

George.Ainslie@va.gov
www.picoeconomics.org

Abstract

Fine examination of our accumulated cultural knowledge is especially helpful in studying the emotions, which are only tangentially accessible to experimental manipulation. Here I use the six properties of emotions that Elster has summarized (as he previously extracted from philosophical, historical, and literary sources) to suggest how they show a need for changes in the science of motivation. The apparent adaptive purpose of emotions lies in their *action tendencies*—what they add to the cold calculation of advantage. Subjectively they stand out by their intrusiveness, the *duration* of which often has a half-life. Emotions each have *valence*, which suggests that they are not only motivating but also motivated, an implication that requires re-examination of how negativity works. Emotions are also experienced as “*triggered*,” but are so malleable that triggering cannot mean simple conditioned reflexiveness. Emotions are not only triggered—or motivated—by *beliefs*, but motivate beliefs in turn, and can be fed back on themselves in a “wildfire” phenomenon. These feedback effects are further evidence against emotions being reflexive.

With regard to the three great revolutions, I argue that enthusiasm differs from romantic love only in its object, and *urgency* comes from the dysphoria of using response suppression for self-control. I can add nothing further to Elster's masterful history.

Key Words

Motivated emotion, motivated belief, natural selection, Pavlovian conditioning, phenomenology, introspective method

¹ This material is the result of work supported with resources and the use of facilities at the Department of Veterans Affairs Medical Center, Coatesville, PA, USA. The opinions expressed are not those of the Department of Veterans Affairs or of the US Government.

Many of the references of which I was author or co-author are downloadable from www.picoeconomics.org

I. Introduction

One of Jon Elster's several projects is to delineate the emotions linguistically, that is, by exploring the properties revealed in common usage as enriched by cameos ranging from proverbs to percepts by the great essayists. This eclectic approach² creates a finer-tuned map than behavioral and neurophysiological methods have so far been able to do (Adolphs & Anderson, 2018; Pessoa 2015). It also covers processes that are commonly called emotions but lack either regular physiological accompaniments or universal expressions. His purpose is to study the "action tendencies" that these processes add to what are assumed to be the baseline or routine (or perhaps "expected") motives governing choice.

Here he applies his method to three political revolutions that were accompanied by strong emotions. In them he finds evidence for a new emotional kind, enthusiasm, which he adds to the perhaps twenty six kinds he has previously described in *Alchemies of the Mind* (1999) and *Explaining Social Behavior* (2015). I will have little to add to his probing history, but will take this opportunity to suggest how his taxonomy supports some modifications to conventional motivational science. Perhaps there is a framework by which these "findings outside the laboratory" (p. 2) can be related to those that came originally from inside.

Before beginning I should say that I don't think his theory of emotion is in conflict with what he calls rational choice theory, which he characterizes as "the imputation to agents of continuous subjective probability distributions" and "imputing steadily declining time discounting over an infinite horizon," among perhaps other examples. He is probably right to imagine that a person called upon to describe her degree of either certainty or time discounting would draw her answer from a small number of qualitative categories. However, that same person makes actual choices according to fine distinctions among continuous values, which she has often been shown experimentally to perceive according to (separate) hyperbolic functions (Green & Myerson, 2004). Defining the exact shape requires parametric experiment (and is still subject to controversy-- Luhmann, 2013, Read & Roelofsma, 2003), but the weighing of quantities from experience is clear to introspection. When I choose whether to take a familiar elevator versus the staircase I'm aware of imagining the stairs and then the faster but unpredictable elevator and then, if the choice is close, the stairs again (the *vicarious trial and error*, or *VTE*, first described by Tolman, 1932; Redish, 2016). In doing so I'm comparing impressions of two efforts at two delays, sampled imperfectly from my recent experience but not forced into categories. This sort of weighing can be seen neuron by neuron in the parietal cortices of monkeys choosing between food rewards (Glimcher, 2009).³ This is not to say that we have experimental data on the motives in emotions, but just that the experiential method does not—and should not—contradict a theory of choice according to prospective discounted reward.

² The term "phenomenology" springs to mind, but as I understand this school it advocates abandoning our collective cultural wisdom—exactly the opposite of Elster's approach. I'll expand on his method at the end.

³ The evaluation of distant outcomes is based on more speculative assumptions (see Rick & Loewenstein, 2008), but is still quantifiable (and still apt to be hyperbolic—Bradshaw, 2019; Green et.al., 2013)

The literature of behavioral science is now ablaze with emotion. However, as Elster points out, the word “emotion” lacks clear boundaries, and the processes that get called emotions are a heterogeneous group. As with pornography, I know it when I see it. Elster explores possible defining features, but finds them not to be strict. Emotions are to be just what are called emotions. Although it might be that people are using the same word to name differing kinds of process, it seems likely that they are recognizing some core feature. I’ll start with that.

II. *Action tendencies*

As Elster points out, a basic property of emotions is their “well-defined action tendencies.” That is, they affect how actions are rewarded. Some emotions are accompanied by activity in distinct parts of the brain; but as has often been pointed out, competing options have to be priced in the same neural currency, or at least in currencies that are commensurable (Cabanac, 1992; Montague & Berns, 2002; Shizgal & Conover, 1996). That is, there must be a mechanism to weigh any behavior against any other behavior that might replace it. The mechanism by which alternative options compete on basis of prospective reward is increasingly well known (Glimcher, 2009), although not specifically for emotions. It is important to note here that reward governs not only deliberate or even conscious processes; it operates even during sleep (Granda & Hammack, 1961; Williams et.al., 1966). Only choices that a person would have the ability to reject can be called deliberate.

The motivation generated by emotions is widely recognized as special. Motivation *tout simple* does not get called emotion. The opportunity to close a million dollar deal may create a strong motive without causing what could be called emotion, at least in a professional investor. So what do emotions add? The human ability to perform precise cost/benefit analyses could have wound up shutting emotions out of any major role: The whole human species could have been shaped genetically to make decisions just on the basis of cold calculation—the outcome that Darwin thought would have been most adaptive (1872/1979). The existence of some people who have much higher thresholds for emotionality than others suggests (Jorgensen et.al., 2007) that such an adaptation would not have taken many generations to shape—that the ingredient processes are on hand to be selected. Since strong emotions persist, we must conclude that they have an ongoing adaptive function, rather than being mere holdovers from a less cognitively developed past. Some possibilities:

- The conventional account is that additional motivation is needed where an individual might under-read a risk, or read it too slowly, for instance in seeing an existential threat to be merely a poorer foraging prospect. That is, in some stereotyped situations it may prove more adaptive to override deliberation. A related pitfall of deliberation is ambivalence. If the odds in favor of attacking a threat are just as favorable as the odds for fleeing, something needs to let one alternative win quickly and stiff-arm the loser.
- A similar factor is that emotion may modify an individual’s effective delay discounting rate. The usefulness of fear in overriding patience is obvious. Conversely, disgust reduces the temporal discounting of nauseating and emetic events, as seen in the bait shyness that develops even when the effects of a poison occur after many hours (Garcia et.al., 1974). The long-lasting social devaluation that is said to be driven by disgust in humans (Schnall et.al., 2008) may or may not provide an evolutionary advantage.

- Sometimes an emotion may supply a game-theoretic strategy that an individual might not learn from the mere contingencies of reward—for instance, to always to punish a trespass, where in a single case it would not seem worth the cost (Nesse, 1998, p. 411).
- Emotion—or mood—may modulate an individual’s overall readiness to act on motives. That is, it may function as a volume knob that sets a general motivational tone. Euphoria entails the lifting of inhibitions, broadening of imagination, and potentially an increase in productiveness (Frederickson, 2013). Its limiting condition is manic psychosis, but there are individuals who ride persistent hypomania into enduring accomplishment (Gartner, 2008). The adaptive function at least for individuals of grief and its corresponding clinical state, depression, are less apparent. Authors have wondered if they deter individuals from pursuing lost causes, perhaps from taking risks when they are vulnerable (Engel & Schmale, 1972), or whether they elicit help from other people (Syme et.al., 2016).
- The most consequential function of emotion may occur where reward maximization diverges from evolutionary adaptiveness. Emotion may be needed to protect an individual’s genome (shared with relatives) at the expense of that particular individual—why else should an organism attack suicidally to defend territory or family? Why even produce children? The threat from such questions to the survival of a genome will be greatest when individuals can weigh long term outcomes, and especially the more they learn to counteract the hyperbolic delay discounting that favors short term rewards. Anger overrides not only ambivalence, but also your concerns for personal safety versus that of your mates and offspring. Love overrides the personal defensiveness that otherwise lets an individual compete with others most effectively, with the result that even foresighted lovers may disregard the costs of bearing children. It has even been suggested that a depressive response to personal failure can lead you to stop competing with your neighbors for resources, in the extreme case by committing suicide (DeCatanzaro, 1980; Syme et.al., 2016).

Group adaptation may also be the role of Elster’s proposed *enthusiasm*. He follows the *Oxford English Dictionary* in defining it as a “rapturous intensity of feeling in favor of a person, principle, or cause...”, but “minus the reference to a person.” The absence of an individual to focus on does not seem to be a crucial difference. Often political enthusiasm is combined with the similar divine love of an individual, as with the Divine Being who leads cultish factions in Hume’s example (p. 10)—surely we saw this in Robespierre.

These theories about adaptiveness do not tell us how emotions function differently from ordinary motivations. To find what kind of motivational configuration is necessary and/or sufficient for emotion, we should examine how Elster describes them. He takes notice of several features in addition to action tendencies:

- [All] Emotions are intrusive. He picks up on the converse of this feature, their limited half-lives, but this highlights the self-renewing tendency that keeps them going until it fades. Often the intrusion entails physiological arousal, but it is sometimes only a motivational configuration that is hard to move past. Envy, say, or pity, pride, or regret, beckon repeatedly as if suggesting action, often when there is no action to take.

- [All] Emotions have valence. That is, whatever motivations for action they may create, their experience itself is either attractive or aversive—or in some cases a combination of both, but the result in those cases is not neutrality.
- [All] Emotions are triggered by beliefs (a “core idea”).
- “A crucial fact is that *beliefs can serve both as causes and effects of emotions*” (his italics), in the latter case when they change the gathering and interpretation of facts. One way that Elster says emotions can shape beliefs is by a feedback effect, e.g. shame about envy leading to its transmutation into anger.
- Emotions can induce urgency (inaction aversion), a preference for action earlier rather than later. Again in terms of reward, such emotions impose a cost for delay over and above a person’s regular discounting of the future.

III. Intrusiveness

When physiological arousal occurs it entails an intrinsic time course, an arc which, once begun, impels you to ride it. In this essay Elster pictures emotions as mental states that are aroused rather than calm, calm meaning that “it does not differ from the baseline physiological state of the organism.” However, in requiring arousal he is narrowing his definition to exclude putative emotions such as envy, pity, pride and regret, which are so called by other writers and by Elster himself in earlier works.

Arousal has dominated the concept since the early 1800s, when moral philosopher Thomas Brown popularized “emotion” as a generic term combining “the passions” and “the affections” (Dixon, 2012).⁴ Among these “vivid feelings” the affections were “milder,” perhaps even intellectual (as Aquinas said), and usually connoted moral good, as if arousal went along with being bad. Negative but non-aroused processes such as envy and pride were seemingly not included, although they should have been part of Brown’s vivid feelings.

When aroused you strive to get something, but are moved by an additional appetite not necessary to the getting. You avoid situations associated with harm and loss, but more than is realistically necessary. You mate for the physical reward, but sometimes feel the need to mate only with a particular person, or indeed do things with that person that are not part of mating and might even involve renouncing the physical act. Your pleasure in a rewarding situation sometimes escalates into euphoria. You respond to a loss with reactions that could not be expected to restore what was lost, and might even get in the way. (Grief reactions might not be well described as arousal, but rather some term like deflation.) Such extra motivation is persistent and falls into a few familiar patterns, which make aroused processes conspicuous and easy to label. These patterns go along with changed body states, which were, until recently, thought to be hardwired subroutines, displayed uniformly across most cultures by involuntary facial expressions (Ekman; 2003, pp. 3-14; Gendron et.al., 2014), analogous to displays in nonhuman animals (Matsumoto et.al., 2008), and accompanied by increasingly well-identified brain activity (Schirmer & Adolphs, 2017; Adolphs & Anderson, 2018). Recent research has found that any named kind of arousal is expressed in many different ways in different subjects (Barrett, 2017; Siegel et.al., 2018), and is accompanied by brain activity that is only roughly localized (Kohn et.al., 2014;

⁴ French authors, including Elster’s favorite Montaigne, used the term earlier, but meaning motivational turbulence in general

Wager et.al., 2015). However, aroused emotions clearly have motivational force, supplying additional action tendencies to the existing motivational landscape.

In addition to these named emotions there is arousal *per se*, perhaps best called excitement, which may be an element of any of the others except grief (and even that in the symptom of agitated depression), and may happen without an identified object. It may happen when watching competitions, or traveling fast, or just by a seemingly random somatic process. In a famous experiment, Schachter & Singer infused subjects with adrenalin, and they reported feeling anger or euphoria depending on fictional cues (1962).

Arousal aside, people often notice that they are moved by a factor beyond the routine incentives created by a circumstance. In his earlier work Elster made a “long overdue distinction” between two meanings of emotion, “occurrent” and “dispositional” (1999a, p. 26)—between actual episodes and a low threshold for having such episodes. That would be a good distinction to keep here, adding that occurrence is often marked by arousal.⁵ That is, there are two kinds of emotional influence, being aroused and simply responding to a provocative configuration of incentives. The same name may often refer to both kinds—you can be both acutely and chronically angry, or fearful, or grief-stricken. Some emotions don’t provoke arousal at all. Even without arousal, anger, say, comes from something that makes a particular provocation hard to ignore. You try to go about your business, but in the midst of your usual incentives there is a distraction, the press of a story that occupies you, or a threat that this new incentive will linger if not taken care of. Or some feature may put a comparison of your own status with someone else’s in the path of your customary self-appraisals, a disturbance that you experience as envy or pity. Or the awareness of a success invites rehearsal, which is more pleasant than the task at hand, but thus goeth before a fall. Or the memory of a past choice tempts you to believe it happened a different way, or to rehearse how it could have, generating regret or guilt. Some groups of seductions get recognized as having similar configurations and are thereafter stereotyped as kinds—those that affront you, those that invite invidious empathy, those that create a wish to undo them, and so on.

The list of durable motivational configurations is probably not finite (Frijda, 2017). Just casting about at random I come up with the incongruity that provokes mirth—aroused, sometimes, as when someone “dies laughing;” the emotions around incomplete knowledge: curiosity, doubt, or suspense; frustration, which, though it classically provokes anger, can just as easily engender boredom. Should we include reminiscence—the lure of memory itself—or only when it takes a particular coloration, as in nostalgia? To the extent that motivational configurations are persistent lures the resulting motives will be experienced as intrusive, even when you are not aroused. Arguably this intrusion, and only this, is what gets a process called an emotion. The other properties in Elster’s list depend on what causes the intrusiveness.

IV. Valence

⁵ People are apt to say “aroused” when they mean just “occurrent.” You may say that your curiosity is aroused, although your physiology has not changed—but less likely that your nostalgia was aroused, and certainly not your boredom.

Elster points out that all emotions have valence. This property is peculiar if, as is usually held, emotions are triggered automatically. It could be argued that their reward value is an epiphenomenon, that they are released like reflexes, without regard to valence. Certainly much emotion is stereotyped, in the sense that it needs specific circumstances to be robust—frustration in the case of anger, threat in the case of fear, attractiveness in the case of love, surprise in the case of joy, loss in the case of grief. Sometimes even the exact circumstance is innately programmed, for instance when fear is provoked by being at a height or seeing a snake-like object. But one of Elster's valuable points is that we feel responsible for our emotions: Recognition of our envy is painful, as is seeing ourselves act in unjustifiable anger. Changing the justification is easier than changing the emotion, so the emotion must be stronger than the supposed trigger, but the implied question remains: if the emotion is not itself motivated, why do we feel responsible for it? And as Elster has noted (1999b), other philosophers have sometimes treated emotions as choices, beginning with Sartre (1939/1948), but they have not reconciled this view with the function of motivation as it is understood in other contexts.

We should look more closely at the nature of valence. Elster follows the generally accepted practice of classifying emotions as having a positive or negative valence, and certainly we are often glad, or not, to find ourselves experiencing a particular emotion. However, his own eclectic method finds evidence that the value of even the arousable emotions is apt to be complex, depending on where you are in their time course and on what options go with them. I have argued elsewhere that all arousable appetites must have a positive—rewarding—component (2001, pp. 65-69; 2017a). This is not just in the trivial sense that all must compete for your attention. We pay to experience anger, fear, and grief in some fictional settings. Although part of this value is in building our appetite for an eventual resolution, one phenomenologist has pointed out that there is pleasure in the fear itself (Hanich, 2011). It is not hard to find examples in daily life of “nursing” these emotions—the sorehead who is always seeking arguments or fights, the ads for how terrifying a roller-coaster is, the bereaved person who rejects chances to rejoin the living. A whole category of emotion is never simply negative.

The case of anger furnishes the best examples of ambiguous value. When another driver cuts me off in traffic, I am conscious of mentally trying out whether to ignore the provocation or to pursue the occasion to get angry, somewhat on the basis of how rewarding the alternative line of thought promises to be. Indeed, Jennifer Lerner (whom Elster cites) has shown that brain activity in anger resembles that in other rewarding situations (e.g., Lerner & Tiedens, 2006). Any arousable emotion has the potential to become addictive—rewarding enough to support repeated choice, but less desirable than alternatives when anticipated at a distance. Arousal itself can be addictive, as seen in the war photographers and emergency room specialists who are drawn to the “adrenaline rush.” Like addictive substances, arousal can habituate, so you are “velocitated,” both needing increased degrees and dreading a subsequent let-down.

Still, people can't usually choose emotions like pastimes, or avoid them like mistakes. Emotion is partly shielded from motives. Not only could we not beat a rhythm with bursts of anger, say, but any deliberate production of an emotion is regarded as false (Frank, 1988). And yet actors and con-persons can do it reliably (Rousseau said they were the same thing—Orwin & Tarco, 1997, 20-44, 274-295), arousal and all. Anyone has some power to induce her own emotions—Ekman listed eight ways (2003, 31ff). The conventional explanation is that you imagine potential

triggers and dwell on them, but that would seem to have the weights in reverse. Emotions have valences and action tendencies, often great ones, whereas imagination itself is facile. When people learn to evoke an emotion habitually—actors are a conspicuous case—they need progressively less trigger scenario, eventually just making a simple choice. The triggering of supposedly involuntary but motivated behaviors is a subject that is especially in need of the eclectic method.

V. *Triggers*

The emotions open up the whole topic of how “involuntary” mental processes are selected, and they are the examples most apt to produce inter-subjective agreement. The main obstacle has been early lab research that led selective factors to be classified as two separate kinds, rather than arrayed along continua of latency and duration. In standard examples where a cue is followed by a response which is followed by a selective process that determines whether the response will be learned, an inability to experimentally manipulate the valence of emotions or other “involuntary” processes such as somatic urges (hunger, pain, urge to excrete...) led to the conventional belief that they are not motivated. For example, to attack a threat is counted as a reward-dependent (or “goal-directed”) behavior, but becoming afraid of it is said to be triggered like a reflex, without regard to what the fear will feel like. The fear response is supposed to be selected by whether it is actually followed by a fear-inducing event. In the language of behavioral psychology the attack is an operant that is selected (*rewarded*) or not by the valence of its consequences. The fear is said to be a respondent, or conditioned response (CR), which is selected by a conditioned stimulus (CS), the effectiveness of which is in turn selected by sometimes leading to an innately powerful unconditioned stimulus (UCS) that is the inborn trigger. This certainly appears to be how many emotions behave. However, there is both laboratory and subjective evidence that the operant/respondent distinction arises from practical experimental limitations, not a difference in the basic selective process.

The notion that respondents are selected by a different selective process than operants is reported to have originated in the translation from Russian of Pavlov’s foundational work (1927). He said that a dog’s salivation to a cue, for instance, was “conditional” on the cue having been paired with food. But in English works “conditional” became “conditioned” (Dinsmoor, 2004), which coincidentally fit the soon-described distinction of reflex-like triggers from operant reinforcers (Skinner, 1930). Once an event had been found to be rewarding an experimenter could use it to teach any kind of operant, but salivation responded only to food. However, if the prospect of food actually selects salivation by the same mechanism as operant reward selects arbitrary behaviors, this could not be tested and so was not considered. It was thought to be enough that salivation could not be trained by other rewards. But salivation might be rewarded endogenously when food is forthcoming, just as anger might be rewarded during frustration or panic when facing catastrophic loss. Behavior *pulled* by reward looks grossly the same as behavior *pushed* by triggers. Where the source of reward can’t be identified the reflex trigger mechanism looks more justified, but the bigger picture holds reasons to back the all-reward model.

Here is where the many subjective puzzles that Elster and others have described can make sense.:

- The triggers for emotions can sometimes be mental states that you drift into, just as mere opportunities can cue urges to smoke (Dar et.al., 2005, 2010) or take a drug (Meyer, 1988), or urinate (Elster, 1999a, p. 227, note 2).

- As Elster has said, drug cravings themselves “are not only cue-dependent and belief-dependent [dependent on the belief that satisfaction is available], but cost-dependent” (Elster, 1999a, p. 73).
- He has borrowed Paul Ekman’s concept of “emotional wildfire” to describe a positive feedback process in which an emotion in effect becomes a stimulus for itself—incubating without further triggering-- again a familiar experience, but a pattern that could not be caused by mere association (see Ainslie, 2010).
- He describes here how enthusiasm can pull itself up by its own bootstraps, and hunger or guilt can be transmuted into—that is, replaced by—anger. Emotions can evolve in subtle ways which would be awkward to account for by changes in triggers. In Elster’s examples, what would be the unconditioned stimulus for the envy that is scripted as anger, or for counter-wishful thinking?

Help from the laboratory is limited but still significant. It is not only the set of emotions that all have a valence. Psychologists Clark Hull (1943) and Neal Miller (1969) long ago noted that all stimuli that can induce conditioning (UCSs) also have valence in the operant sense. Some experimental theorists have proposed that operant and respondent selectors are indeed the same (Pear & Eldridge, 1984; Donahoe et.al., 1993). My own experiments have shown that a conditioned response—increased heart rate to a cue predicting shock—can be reversed by operant training, implying that the selective processes are at least comparable (Ainslie & Engel, 1974). I have developed this argument at more length elsewhere (2010), including the key factor that has always made an unmotivated pairing process such as conditioning appear necessary: the way that individuals are coerced to participate in negatively valued experiences—not just to rage but to panic, to grieve, and to be disgusted.

New reports about salience bear on the lure of aversive experience, so I will briefly sketch my argument: The imperative quality of unwanted (or disliked—see Berridge, 2009) experiences is often described as “salience,” (Berridge, 2007; Hird et.al., 2018), but since salience often has to compete with reward to attract attention it must be on the same motivational dimension as reward, and in the positive direction. Berridge originally called salience “non-hedonic reward” (2003). Recently “utility” has been used to distinguish the hedonic kind (Hird et.al., 2018), but the implication of a common marketplace that selects the direction of attention has been ignored (however, see Bromberg-Martin et.al., 2010; Ainslie, 2009a). “Salience” was originally a cognitive term, for a property of information, so it might seem motivationally neutral; but in that case the attractive power of salient perceptions remains to be accounted for. Better to model salience as the offer of short-term reward, like that for an itch but with its reward even more closely mixed with dysphoria.⁶ Of course choices with such short latencies as attention and emotional response are usually not deliberate; but people overcome fears of heights and snakes, more or less, and even panic is experienced not as a reflex but as an urge, which people who repeatedly face dire situations learn to resist. Therapy helps people learn how to withhold panic (Clum, 1989) and the emotional (protopathic) component of pain (Sternbach, 1968; Hilgard & Hilgard, 1994, 86-165).when facing the supposedly conditioned stimuli for them. It is not just

⁶ Since “reward” commonly connotes pleasure, rather than its basic function—that which induces repetition of what it follows-- we offend the ear less if we call the influence of a short term reward such as salience an *urge*, rather than, say, an appetite, a hunger, or a desire. Of course there are urges for pleasures as well-- the term is inclusive.

stoics who learn to be a-pathetic. Triggers for emotions may sometimes be “offers you can’t refuse,” but they are still offers rather than reflex-releasers, and people sometimes refuse them.

Assuming thus that even aversive emotions are based on a rewarding component, it is reasonable to propose a motivational model in which learned pathways to reward are not just tools that lie about waiting to be called on, but active processes—I’ve used the term *interests*—that forage for reward, perhaps like domestic pets that stay alert to when a route to reward might be opening up. Where the environment presents you with concrete threats and resources this model will be hard to tell apart from the conventional model in which reward depends on realism—that is, in which behaviors are shaped by *secondary rewards* that are given power to the extent that they predict external *primary rewards*. However, where primary rewards are remote or are matters of interpretation, choice depends on imagined—*endogenous*—rewards, which win not by realism but by the reward promised by their occurrence itself. The extra reward called up by emotion is a major example of the endogenous category. Of course this model opens up the question of how endogenous reward is constrained, if not by hard facts,

VI. Beliefs

Elster’s most penetrating observation is the reciprocal action of emotions and beliefs. Functionally, belief is the perception of constraint on choice. This constraint may be trivial and *ad hoc*, such as an expectation that I will not have time to go through a traffic light before it turns red. Even such a trivial perception may be a case of belief *causing* emotion—a pause in my joy of driving, a blip of fear that I will break the law—or belief *caused by* emotion—incidental euphoria or anger giving me a more optimistic expectation, sadness a pessimistic one. Such mindless expectations are easily explained by conventional motivational science. The harder case are constraints that are at least partly composed of what the beliefs themselves imply. “Does driving late into the yellow period amount to running the red light?” “If I run the red light, will it bother me that the anonymous strangers who see it disapprove of me?”, or even, at an isolated intersection, “If I run the red light this time, will I have to spend attention deciding whether to run the next isolated one, and the next...?” These beliefs are matters not just of estimation, but of interpretation, and thus even more subject to motivation. The same euphoria or anger may affect such interpretations, which, unlike simple expectations, set precedents. To the extent that I excuse my choice, not let expected social opinion bother me, or look for a line distinguishing “isolated” traffic lights from others, I have shifted my criteria for how much to take chances, what weight to give others’ opinions, or whether to respond to regulations on a case basis. To the extent that this process itself seems to entail dangerous uncertainty, I may perceive the incentives to avoid it as *facts*: Failing to brake on yellow *is* dangerous, others’ opinions *are* crucial, traffic rules *transcend* the value of outcomes.

In a model without the mechanism of conditioned responses, endogenous reward is a fiat currency, issued at will but subject to the risk of inflation-- dissipation of the appetite for this reward when the same kind is issued too often. Arousable emotions and other endogenously rewarded processes are selected when a cue predicts sufficient appetite for them and good pacing to sustain them. Such cues *occasion* these processes rather than unleashing them like reflexes.⁷

⁷ The emotional effectiveness of singular occasions may be experienced as a kind of factuality, more or less confounded with the factuality that comes from physical observation. In the most

For instance, anger is not the most satisfying of emotions, but if a peasant in Elster's account did not have good occasions for joy or pride, and indeed had urges for envy, anger might have been the best-paying activity. However, anger would just support transient daydreams unless a few singular occasion for it stood out from all others.⁸ The greatest singularity usually comes from hard facts, but the scenario from a set of coincidences behind a paranoid rumor, for instance about brigands destroying crops, could perform the same service. Once a single delusion was established, a change of cast from brigands to landlord nobles did not impair it. A mob of neighbors embracing the rumor might put it over his threshold for belief-- and he over theirs, Ouija fashion-- and it might not matter that he had doubts about the evidentiary value of coincidences and mobs; the occasion would still stand out. Here is a positive feedback cycle of motivated belief. When the belief is also a personal rule—"resisting oppression is patriotic;" "landlords can't be trusted"-- the feedback effect is stronger, for disbelief would represent a failed test case (Ainslie, 2009b, 2017).

The strong motivation that supports enthusiasm, like romantic love, seems to come from occasions to throw caution to the winds. People have learned to be habitually cautious with each other, and test each possible cooperation against the many scenarios in which it could go wrong. When a unique combination of factors seems to make this caution unnecessary, the heady sense of permission for all your pent-up wishes is what Elster reports for enthusiasts, and the poets report for lovers. This perception of singularity must begin somewhere, of course, but need not start with the force of a trigger in the Pavlovian sense, which is what Elster is puzzled about.

VII. Urgency

conspicuous cases, remembered events are experienced again on their anniversaries, especially when the anniversary is a round number; original works of art are felt to be more "real" than exact copies; and placebos are effective in proportion to the expensiveness of the ingredients or the prestige of the healer. Even realistic beliefs get additional value by serving as occasions for emotional reward, as in the "drug effect" of money (Lea & Webley 2006). Conversely, faced with unwelcome urges such as hypochondria, phobic anxiety, or a sense of being dirty, a person searches for a favorable interpretation of the situation – whether she can feel well, or safe, or clean. This interpretation cannot be arbitrary; wishes have little impact. She must choose her belief on the basis of "facts" that she discerns in events beyond her control – a pill given by a doctor, a lucky charm or safety signal, or a "scientific" disinfectant. The belief may even become stabilized as a personal rule: in effect, "I will not give in to panic or disgust when this signal is present."

⁸ Here the important aspect is that imagination ad lib exhausts itself in premature payoffs. When one occasion for reward is as good as another, they will replace each other randomly, and the imagining will have the quality of a daydream. Conversely, if there is a single, relatively rare occasion that stands out from the others, it will make the corresponding imagination robust. The experience of such *singularity* may be much like that of having solved a puzzle or detected a fact of nature. The occasion in question will stand out from the common ruck of imaginings just as a fact stands out from a fantasy.

Elster has adopted this term to name a specific source of the tendency to act sooner rather than later. It needs to be understood as narrower than the common usage, which includes extreme impatience, such as the urge to pull your hand away from a hot radiator. It looks to me like it names the effort of using mental suppression to defer action. I have argued that what gets called willpower may be either resolve—the product of intertemporal bargaining—or suppression—moment-to-moment vigilance against a specific action (Ainslie, 2021). Suppression is the more primitive, empathically evident in a dog awaiting its master's fetch signal, but it is the one modeled in most laboratory tests of willpower (Friese et al., 2018). Suppression is costly because it ties up cognitive resources (Kurzban et al., 2013). The supposed depletion of willpower by tedious tasks could be said to come from cumulative urgency. I would say that the anguish of Kamikaze pilots came from this process raised to a power.

I don't know if Elster would count the motivation for sooner action that comes from a very different mechanism, the wish to harvest emotional reward before the emotion wanes. People are often quite aware of emotional half-lives, and so may realistically seek, for instance, to retaliate before they are no longer mad, or satisfy curiosity about a question before they forget it. In a nearly analogous situation, I have often avoided dissipating sleepiness with caffeine so I can enjoy a nap.

VIII. Is this science?

This article helps define a niche in motivational science that Elster has been developing for most of his career. He has previously cast a net over the behavioral sciences (in which I include philosophy), the humanities, and history, including voices from the past—not just the canonical philosophers but also Emily Dickinson, John Donne, Stendhal, Montesquieu, Tocqueville, and especially Montaigne--mining introspections and presenting them in a way that enables inter-reader agreement. The subjectivist method has been called phenomenology, but as I understand it, the teaching of the original phenomenologists was to blank out generations of cultural assumptions and start with raw introspection (Crotty, 1996). Elster's approach is quite the opposite—examination of cultural statements in the belief that they often had a basis in personal observation

The most important mental processes are subtle, even recursive—the self observing the self—and at most tangentially accessible to controlled experiment. But many people have observed these processes, in themselves, and many have written them down. Some behavioral scientists sniff at experiential evidence as "folk psychology." However, while common sense is suggestible at best and, as theory, almost always inconsistent and ad hoc, it is by far the largest body of human observations. Elster wrote in *Alchemies of the Mind* (1999b), "With respect to an important subset of the emotions we can learn more from moralists, novelists, and playwrights than from the cumulative findings of scientific psychology... Prescientific insights into the emotions are not simply superseded by modern psychology in the way natural philosophy has been superseded by physics. Some men and women of the past have been superb students of human nature... There is no reason why one century out of twenty-five should have a privilege in wisdom and understanding" (p. 50). In the conclusions of *Explaining Social Behavior* he names names and gives details (2015), the gist of which is that it is arrogant for experimentalists to dismiss observations that have not undergone their own certainly imperfect process of

sanitization. Elster's longstanding criticism has recently been validated by the meta-analytic finding that controlled experiments on cognitive phenomena have generally shown small effects and often failed of replication (Friese et.al., 2018; Hagger et.al., 2016; Simmons et.al., 2011).

Behavioral science in the twentieth century mostly disregarded internally conflicting or inconsistent motives, seemingly on the assumption that, once contingencies of reward were clear, a person's choice would be obvious. However, this assumption was contradicted by the body of experiments that became known as prospect theory (Tversky & Kahneman, 1981). These were simple procedures that usually consisted of tallies of paired introspections. Subjects did not choose between actual medical treatments that could save 400 lives or lose 200, but merely expressed hypothetical preferences between them. The situations were probably thought up in much the same way as Elster's emotional triggers, but were then presented to convenience samples of subjects—mostly American students. The empirical surveys were useful safeguards, of course, and sometimes have been revealed cultural differences in subject populations (Rieger et.al., 2011), but mostly could have been tested in the reader's own introspection.⁹ Elster's examples in his current proposal might well be testable in the manner of Kahneman and Tversky ("If you wanted revenge for a wrong, would you rather have it now or in a week..."), but it would be subject to many of the very distortions that he is writing about ("Revenge? Me?"), as well as requiring quite bit of self-awareness to begin with. A dedicated experimentalist might say that the facts are thus un-knowable, but this would be to discard the high degree of inter-subjective agreement I expect that his illustrations elicit—although an objection of Western cultural specificity could stand, just as in the case of game theory (Henrich et.al., 2010). I don't see how Elster's eclectic method differs in kind from the linguist's study of a native speaker (but see Love & Ansaldo, 2010). It is a pillar of motivational science.

Declaration of Interest

I have no conflicts of interest

References

Adolphs, R. & Anderson, D.J. (2018) *The Neuroscience of Emotion: A New Synthesis*. Princeton U.

Ainslie, G. (2001) *Breakdown of Will*. New York, Cambridge U.

Ainslie, G. (2009a) Pleasure and aversion: Challenging the conventional dichotomy. *Inquiry* 52 (4), 357-377.

⁹ Some of these demonstrations suffered from their categorical nature, often because subjects obeyed what they took to be the expectations of public surveys, or did not relate their terms to real life situations (Gigerenzer et.al., 2012) It would be interesting to know what comparisons, if any, were tried out on subjects but failed to evoke differences.

- Ainslie, G. (2009b) Recursive self-prediction in self-control and its failure. In Till Gruene-Yanoff and Sven Ove Hansson, eds, *Preference Change: Approaches from Philosophy, Economics, and Psychology*. Springer, pp. 139-158.
- Ainslie, George (2010) The core process in addictions and other impulses: Hyperbolic discounting versus conditioning and cognitive framing. In D. Ross, H. Kincaid, D. Spurrett, and P. Collins, eds., *What Is Addiction?* MIT, pp. 211-245.
- Ainslie, G. (2017) *De gustibus disputare*: Hyperbolic delay discounting integrates five approaches to choice. *Journal of Economic Methodology* 24(2), 166-189.
<http://dx.doi.org/10.1080/1350178X.2017.1309748>
- Ainslie, G. (2021) Willpower With and Without Effort *Behavioral and Brain Sciences* 44
- Ainslie, G. & Engel, B.T. (1974) Alteration of classically conditioned heart rate by operant reinforcement in monkeys. *Journal of Comparative and Physiological Psychology* 87, 373-383.
- Barrett, L F. (2017). *How emotions are made: The secret life of the brain*. Houghton Mifflin Harcourt.[Re Elster Emot in Sw lib]
- Berridge, K.C. (2003) Pleasures of the brain. *Brain and Cognition* 52, 106-128. [have]
- Berridge, K.C. (2007) The debate over dopamine's role in reward: The case for incentive salience. *Psychopharmacology* 191, 391-431.
- Berridge, K.C. (2009) Wanting and liking: Observations from the neuroscience and psychology laboratory. *Inquiry* 52, 378-398.
- Bradshaw, C.M. (2019). In search of a definition of reinforcer value: Some successes and failures of the multiplicative hyperbolic model. *Behavioural Processes* 167, 103884.
- Bromberg-Martin, E.S., Matsumoto, M., & Hikosaka, O. (2010). Dopamine in motivational control: rewarding, aversive, and alerting. *Neuron*, 68(5), 815-834.
- Cabanac, M. (1992) Pleasure: The common currency. *Journal of Theoretical Biology* 155, 173-200.
- Skinner, B.F. (1930) On the conditions of elicitation of certain eating reflexes. *Proceedings of the National Academy of Sciences* 16, 433-438.
- Clum, G A. (1989) Psychological interventions vs. drugs in the treatment of panic. *Behavior Therapy* 20, 429-457.
- Crotty, M., (1996). Doing phenomenology. In *Qualitative Research Practice in Adult Education* eds. Lovell Publishing. Pp. 272-282
- Darwin, C. (1872/1979) *The Expressions of Emotions in Man and Animals*, London: Julian Friedman Publishers.

- DeCatanzaro, D. (1980). Human suicide: A biological perspective. *Behavioral and Brain Sciences*, 3(2), 265-272.
- Dixon, T. (2012). "Emotion": The history of a keyword in crisis. *Emotion Review*, 4(4), 338-344.
- Dinsmoor, J.A. (2004) The etymology of basic concepts in the experimental analysis of behavior. *Journal of the Experimental Analysis of Behavior* 82, 311-316.
- Donahoe, J.W., Burgos, J.E. & Palmer, D.C. (1993) A selectionist approach to reinforcement. *Journal of the Experimental Analysis of Behavior* 60, 17-40.
- Ekman, P. (2003) *Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life*. Times Books.
- Elster, Jon (1999a) *Strong Feelings: Emotion, Addiction, and Human Behavior*. Cambridge, MA, MIT.
- Elster, J. (1999b) *Alchemies of the Mind: Rationality and the Emotions*. Cambridge U. Press.
- Elster, J. (2015) *Explaining Social Behavior: More Nuts and Bolts for the Social Sciences*. Cambridge U. Press.
- Engel, G. & Schmale, A.H. (1972). Conservation-withdrawal. In: Physiology, Emotion and Psychosomatic Illness ed. R. Porter & J. Knight. In *Ciba Foundation Symposium* (Vol. 8, pp. 57-85).
- Frank, R.H. (1988) *Passions Within Reason*. New York: W.W. Norton and Company.
- Fredrickson, B.L. (2013). Positive emotions broaden and build. In *Advances in experimental social psychology* (Vol. 47, pp. 1-53). Academic Press.
- Friese, M., Loschelder, D.D., Gieseler, K., Frankenbach, J., & Inzlicht, M. (2018). Is ego depletion real? An analysis of arguments. *Personality and Social Psychology Review*, 1088868318762183.
- Frijda, N.H. (2017). *The Laws of Emotion*. Psychology Press.
- Garcia, J., Hankins, W. & Rusiniak, K. (1974) Behavioral regulation on the milieu interne in man and rat. *Science* 185, 824-831.
- Gartner, J.D. (2008). *The hypomanic edge: the link between (a little) craziness and (a lot of) success in America*. Simon and Schuster.
- Gendron, M., Roberson, D., van der Vyver, J.M., & Barrett, L.F. (2014). Perceptions of emotion from facial expressions are not culturally universal: evidence from a remote culture. *Emotion*, 14(2), 251.

- Gigerenzer, G., Fiedler, K., & Olsson, H. (2012). Rethinking cognitive biases as environmental consequences. In P.M. Todd, G. Gigerenzer, & the ABC Research Group (Eds.). *Ecological Rationality: Intelligence in the World* (pp. 80–110). New York: Oxford University Press.
- Glimcher, P.W. (2009) Choice: towards a standard back pocket model. In P.W. Glimcher, C. Camerer, Russell, A.P., & Fehr, E., eds., *Neuroeconomics: Decision Making and the Brain*. Elsevier, pp. 503 – 521.
- Granda, A.M. and Hammack, J.T. (1961) Operant behavior during sleep. *Science* 133, 1485-1486.
- Green, L. Myerson, J. (2004) A discounting framework for choice with delayed and probabilistic rewards. *Psychological Bulletin* 130, 769-792.
- Green, L., Myerson, J., Oliveira, L., & Chang, S.E. (2013). Delay discounting of monetary rewards over a wide range of amounts. *Journal of the Experimental Analysis of Behavior*, 100(3), 269-281.
- Hagger, M.S., Chatzisarantis, N.L., Alberts, H., Anggono, C.O., Batailler, C., Birt, A.R., ... & Calvillo, D.P. (2016). A multilab preregistered replication of the ego-depletion effect. *Perspectives on Psychological Science*, 11(4), 546-573.
- Hanich, J. (2011). *Cinematic Emotion in Horror Films and Thrillers: The Aesthetic Paradox of Pleasurable Fear*. Routledge
- Henrich, J., Heine, S.J., & Norenzayan, A. (2010). The weirdest people in the world?. *Behavioral and Brain Sciences*, 33(2-3), 61-83.
- Hilgard, E.R. & Hilgard, J.R. (1994) *Hypnosis in the Relief of Pain, Revised Edition*. New York, Brunner/Mazel.
- Hird, E.J., El-Deredy, W., Jones, A., & Talmi, D. (2018). Temporal dissociation of salience and prediction error responses to appetitive and aversive taste. *Psychophysiology*, 55(2), e12976.
- Hull, C.L. (1943). *Principles of Behavior* (Vol. 422). New York: Appleton-Century-Crofts.
- Jørgensen, M. M., Zachariae, R., Skytthe, A., & Kyvik, K. (2007). Genetic and environmental factors in alexithymia: a population-based study of 8,785 Danish twin pairs. *Psychotherapy and Psychosomatics*, 76(6), 369-375.
- Kohn, N., Eickhoff, S.B., Scheller, M., Laird, A.R., Fox, P.T., & Habel, U. (2014). Neural network of cognitive emotion regulation—An ALE meta-analysis and MACM analysis. *NeuroImage*, 87, 345–355.

- Kurzban, R., Duckworth, A., Kable, J. W., & Myers, J. (2013) An opportunity cost model of subjective effort and task performance. *Behavioral and Brain Sciences* 36, 661-726.
- Lea, S.E.G. & Webley, P. (2006) Money as tool, money as drug: The biological psychology of a strong incentive. *Behavioral and Brain Sciences* 29, 161-209.
- Lerner, J.S., & Tiedens, L.Z. (2006) Portrait of the angry decision maker: How appraisal tendencies shape anger's influence on cognition. *Journal of Behavioral Decision Making* 19, 115-137.
- Love, N., & Ansaldo, U. (2010). The native speaker and the mother tongue. *Language Sciences*, 32(6), 589-593.
- Luhmann, C.. (2013). Discounting of delayed rewards is not hyperbolic. *Journal of experimental psychology: learning, memory, and cognition*, 39(4), 1274.
- Matsumoto, D., Keltner, D., Shiota, M.N., O'Sullivan, M. & Frank, M. (2008). Facial expressions of emotion. *Handbook of emotions*, 3, 211-234.
- Meyer, R.E. (1988) Conditioning phenomena and the problem of relapse in opioid addicts and alcoholics. In Ray, B. (ed.) *Learning Factors in Substance Abuse* NIDA Research Monograph series 84, 161-179. NIDA.
- Miller, Neal (1969) Learning of visceral and glandular responses. *Science* 163, 434-445.
- Montague, P.R. & Berns, G.S. (2002). Neural economics and the biological substrates of valuation. *Neuron*, 36, 265-284.
- Nesse, R.M. (1998). Emotional disorders in evolutionary perspective. *British Journal of Medical Psychology*, 71,397– 416.
- Orwin, C., & Tarcov, N., eds. (1997) *The Legacy of Rousseau*. U. Chicago.
- Pavlov, I.P. (1927) *Conditioned Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex*. (G.V. Anrep, trans.). Oxford.
- Pear, J. J. & Eldridge, G.D. (1984) The operant-respondent distinction: Future directions *Journal of the Experimental Analysis of Behavior* 42, 453-467.
- Pessoa, L. (2015) Précis of *The Cognitive-Emotional Brain*. *Behavioral and Brain Sciences* 38, 1-66 doi:10.1017/S0140525X14000120, e71.
- Read, D. & Roelofsma, P.H.M.P (2003) Subadditive versus hyperbolic discounting: A comparison of choice and matching. *Organizational Behavior and Human Decision Processes* 91, 140-153.
- Redish, A.D. (2016). Vicarious trial and error. *Nature Reviews Neuroscience*, 17(3), 147-159. doi: 10.1038/nrn.2015.30

- Rick, S. & Loewenstein, G. (2008) Intangibility in intertemporal choice. *Philosophical Transactions of the Royal Society B* 363, 3813-3824.
- Rieger, M O., Wang, M., & Hens, T. (2011). Prospect theory around the world. *NHH Dept. of Finance & Management Science Discussion Paper*, (2011/19).
- Sartre, J.-P. (1939/1948) *The Emotions: Sketch of a Theory* (B. Frechtman, trans.) Philosophical Library.
- Schachter, S. & Singer, J. (1962) Cognitive, social, and physiological determinants of emotional state. *Psychological Review* 69, 379-399.
- Schirmer, A., & Adolphs, R. (2017). Emotion perception from face, voice, and touch: comparisons and convergence. *Trends in cognitive sciences*, 21(3), 216-228.
- Schnall, S., Haidt, J., Clore, G.L., & Jordan, A.H. (2008). Disgust as embodied moral judgment. *Personality and Social Psychology bulletin*, 34(8), 1096-1109.
- Shizgal, Peter, & Conover, Kent (1996) On the neural computation of utility. *Current Directions in Psychological Science* 5, 37-43.
- Siegel, E H., Sands, M.K., Van den Noortgate, W., Condon, P., Chang, Y., Dy, J., ... & Barrett, L.F. (2018). Emotion fingerprints or emotion populations? A meta-analytic investigation of autonomic features of emotion categories. *Psychological bulletin*, 144(4), 343.[Physical signs not specific to emotions.
- Simmons, J.P., Nelson, L.D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological science*, 22(11), 1359-1366.
- Skinner, B.F. (1930) On the conditions of elicitation of certain eating reflexes. *Proceedings of the National Academy of Sciences* 16, 433-438.
- Sternbach, R.A. (1968) *Pain: A Psychophysiological Analysis*. New York: Academic Press.
- Syme, K.L., Garfield, Z.H., & Hagen, E.H. (2016). Testing the bargaining vs. inclusive fitness models of suicidal behavior against the ethnographic record. *Evolution and Human Behavior*, 37(3), 179-192.
- Tolman, E. (1932) *Purposive Behavior in Animals and Man*. New York: Century.
- Tversky, A. & Kahneman, D. (1981) Framing decisions and the psychology of choice. *Science* 211, 453-458.
- Wager, T.D., Kang, J., Johnson, T.D., Nichols, T.E., Satpute, A.B., & Barrett, L.F. (2015). A Bayesian model of category-specific emotional brain responses. *PLoS computational biology*, 11(4), e1004066

Williams, H.L., Morlock Jr, H.C., & Morlock, J.V. (1966). Instrumental behavior during sleep. *Psychophysiology*, 2(3), 208-216.