Advantages of bias and prejudice: an exploration of their neurocognitive templates

A. Tobena[a,*], I. Marks[b], R. Dar[c]

[a]*Autonomous University of Barcelona, School of Medicine, Unit of Medical Psychology, Campus of Bellaterra, 08193 Bellaterra, Barcelona, Spain*
[b]*Institute of Psychiatry and Maudsley Hospital, London SE5 8AF, UK*
[c]*Department of Psychology, Tel Aviv University, Tel Aviv, Israel*

Abstract

Bias is common in mental-processing tasks as diverse as target recognition, heuristic estimation and social judgment. This paper holds that cognitive biases stem from the covert operation of neural modules, which evolved to subserve adaptive behavior. Such modules can be innate or forged early in development. Research shows links between (i) biases in cognitive tasks and (ii) neural devices, which may mediate them. Evidence is included from biases that arise spontaneously in artificial neural networks during recognition/decision tasks. Two linked propositions follow. First, there are continuities in biasing strategies across different levels of cognitive processing. Second, a proclivity for stereotyping and prejudice depends on the biased functions of lower-level neural modules that promote adaptations to social environments. The propositions rest on evidence of biological preparedness for stereotyping and of deficits in social judgment in patients with neurological lesions. To test such claims, research studies are suggested at the boundary of cognitive neuroscience and social psychology. Advantages of bias and prejudice as evolved tools may include their: (1) speeding of scrutiny and improving of target detection in changing or uncertain situations; (2) aiding of a rapid choice of practical short-term rather than optimal longer term plans; (3) allowing appraisal of a workable world by creating fairly stable categories; (4) motivating of exploration and completion of problem-solving which might otherwise be abandoned too early. The biological priming of social biases need not mean that they are immutable; understanding them could lead to better ways of controlling them. © 1999 Elsevier Science Ltd. All rights reserved.

Keywords: Bias; Neurocognitive templates; Social prejudice

The human understanding, once it has adopted opinions, either because they were already accepted and believed, or because it likes them, draws everything else to support and agree with them. And though it may meet a greater number and weight of contrary instances, it will, with great and harmful prejudice, ignore or condemn or exclude them by introducing some distinction, in order that the authority of those earlier assumptions may remain intact and unharmed. So it was a good answer made by that man who [according to Cicero 106–43 BC], on being shown a picture hanging in a temple of those [who] through having taken their vows, had escaped shipwreck, was asked whether he did not now recognize the power of the gods. He asked in turn: "But where are the pictures of those who perished after taking their vows?" (Francis Bacon, 1620, The Novum Organum).

1. Introduction

How many facts it will take to change a cherished view depends on whether they confirm or refute it. The most trivial fact is seized to confirm pre-existing notions, but contrary evidence must usually be massive to overthrow a strong prejudice and even then may fail to do so. Biased thinking is natural in laymen and in eminent philosophers and scientists. Charles Darwin [25] noted that he was more likely to forget the observations or thoughts that opposed his hypotheses than those that supported them. He guarded against this bias by making a habit of writing a memo of confuting facts or ideas as soon as they came to mind. However, he also pointed out that excessive skepticism can deter people from pursuing a potentially useful train of enquiry. The bias noted by Darwin is a confirmation bias that has been studied extensively [60]. People tend to search for evidence to confirm their favored hypotheses, and avoid or ignore potentially disconfirming evidence.

The tendency to disregard inconvenient facts reflects a pervasive inclination to prime certain features, paths or strategies in cognitive processing. Such inclinations result in

* Corresponding author. Tel.: + 935-811438; fax: + 935-811435.
E-mail address: ikpia@cc.uab.es (A. Tobena)

particular outcomes that may be distorted—when measured by external standards—but are well suited to attain an organism's goals. By bias is meant a systematic preference, which primes either the selection of targets or the recruitment of operational modes, that results in weighted options in cognitive routines. Such weighting usually produces efficient performance but can also lead to false observations or deductions. Bias in a particular task is shown when performance departs systematically from results yielded by consistent (external) measures or by the use of formal rules of objective reasoning.

Visual illusions which stubbornly resist external disconfirmation were the earliest evidence that the brain captures the world through biased processing, but after a century of neurophsychological research different types of bias have been described in the normal functioning of many cognitive tasks. Attention, perception and recall are usually directed towards prominent or selected elements of a given scenario rather than to all the relevant targets or to random sets of those targets [18,22,29,36], and the same is true of the processes of comparing and assessing in some reasoning tasks [79,87,97,108]. Bias is so common in mental functioning that hardly any area of human cognition is devoid of it [87]. Research over the last three decades has concentrated on biases in basic processes, mediating, learning and reasoning, and in social cognition too [6,79,87,97,108].

Though bias can lead to error, flagrant errors usually stem less from bias than from fatigue, distraction, routine, stress, handicap, lesions, drugs, lack of skill and sensorimotor limits, and are easily recognized as such. In contrast, systematic bias is shown by alert well-functioning subjects and acknowledged, if at all, only with surprise or disbelief. Most errors are obviously maladaptive whereas biases are usually adaptive byproducts of cognitive routines that evolved to increase fitness.

This article describes systematic biases in tasks as diverse as target recognition, heuristic estimation and social cognition, and shows how they aid problem solving in various domains. We assume that cognitive biases reflect the silent operation of neural modules that enhance adaptive behavior. 'Neural' denotes dedicated systems in the brain [18,62], and 'silent' means that the working of those systems is not immediately accessible to conscious scrutiny [107]. We thus propose that cognitive biases arise from properties of specialized neural modules that work automatically. To support this idea, we discuss biases in cognitive tasks for which plausible neural modules have been proposed, including biases that arise spontaneously in artificial neural networks performing recognition/decision tasks. Thereafter two propositions follow, that: (1) there are continuities in biasing strategies across different levels of cognitive processing; (2) social stereotyping and prejudice are due to the biased function of neurocognitive modules that usually promote adaptive behavior in social environments. In other words, social stereotyping and prejudice may depend on "prepared" properties of human neural machinery.

Our analytic template is the basic process of categorization. It affects how we construe the world in biased terms from object perception to social stereotyping. We will show that stereotypes appear early in infancy and that focal neurological lesions disrupt typical social biases. We outline research programs that might link social psychology to cognitive neuroscience and uncover plausible biological roots of prejudice. Finally, we propose evolutionary advantages that underpin biased thinking and discuss ideological traps that bedevil the area.

2. Biased cognition

2.1. From biased perception to biased judgement

Systematic biases were first detected in perceptual and learning tasks, but cognitive biases are currently seen as distortions of reasoning. Sound judgment depends, however, on neurocognitive operations involving many processes (e.g. veridical perception, focused attention, short-term storage, association, retrieval, and ability to sequence, compare and combine objects and frames) whose substrates are neural systems with specific constraints and operational modes [5,17,36,88]. Such constraints may contribute to common biases in human thinking.

In visual tasks, the same extraction and priming of relevant features that facilitate veridical perception [45,72,84] lead to illusions—systematic distortions of external reality [46]. Some illusions arise out of evolved rules for fast and efficient grouping of visual objects in the primary neocortex [47,103,109]. Thus, the neural rules that usually yield efficient discrimination can produce errors, ambiguities and paradoxical perceptions. Moreover, while some illusions require awareness of impending stimuli [106] others do not, depending instead on stages of processing which are outside awareness [23,51,52].

Similar constraints that channel relevant features of incoming stimuli also occur in nonvisual perceptual systems. Listeners overestimate the change in level of rising level tones relative to equivalent falling level tones [78]. In a natural environment, this overestimation could be advantageous because rising intensity commonly signals movement of the sound source towards an organism. The bias to hear rising rather than falling tones is even stronger at higher sound intensities, suggesting that increase in loudness becomes even more crucial when a sound source is close or the sound is loud. Speech perception requires high-speed detection of chunks of sound and tonal inflections in the complex stream of sound that speech consists of [3,86]. This restricted mapping of incoming sounds is essential for comprehension but can lead to common misunderstandings and to systematic bias. For instance, when people speak to one another face-to-face, speech perception is further enhanced by a cross-modal bias to match sounds

with lip movements. Though such "forced" matching may lead to illusions (the ventriloquist effect), the illusion itself improves the comprehension of a stream of speech out of several competing ones by enhancing selective listening [28].

Perceptual biases may be boosted by learning biases. Automatic forms of associative learning proceed along pathways, which are not equipotential [20] despite the great flexibility and plasticity of the nervous system. Associations central to survival and reproduction are learned quickly after only one pairing [37,91] or contextual links outside awareness [19]. Yet, other crucial relationships between events are learned very slowly or not at all [71]. Novelty, danger, reward and emotionally charged stimuli are primed paths for biased learning that can often occur outside awareness. Learned reactions to threatening or friendly faces appear automatically even when the faces are presented in masking paradigms that preclude conscious awareness of the faces [81]. If subjects were shown a frightened face for less than 40 ms and immediately afterwards were shown an expressionless mask for far longer, they reported seeing the mask but not the face; despite that, the masked face outside awareness elicited overt emotional responses and amygadala activation on neuroimaging [75,112].

Moreover, in recent studies of more complex learning, the acquisition of initial linguistic skills depended on restricted ways of mapping sounds. For example, 2–5-month-old babies attended more to sounds typical of their native language [64], and 8-month-old infants recognized iterated segments within stretches of nonsensical "wording" sounds [102]. Infants' attention thus seems to be automatically guided towards specific sound features and profiles that will later form templates for perceived "word" categories. Once linguistic categories are established their enhancement appears in the brain. When Estonian and Finnish adults who speak closely related languages heard sounds consisting of a phonetic segment embedding a vowel characteristic of one language but not the other, enhanced evoked potential and magneto-encephalographic responses appeared only in those adults whose language contained that vowel [77]. It exemplifies the basis of the Pygmalion phenomenon whereby arbitrary speech sounds become hallmarks distinguishing one social group from one another.

In summary, these perceptual and learning abilities depend on the operation of neurocognitive systems that use characteristic strategies (choose salient features, neglect monotonous or irrelevant backgrounds, enhance contrasts, segment patterns or boundaries, generalize across similarities, iterate through familiar paths etc.), that lead to efficient performance but also to bias. Similar strategies are also found in judgment tasks.

2.2. Fallacious but adaptive heuristics

In assessing probabilities, comparing odds, or judging from new or uncertain data, people tend to make selective assumptions and to scan only a few of the relevant components of a given problem. Solid research shows that this can lead to departures from objective reasoning [6,33,87,97,108], and to self-deception [74]. People tend to be unduly pessimistic about unfamiliar as opposed to familiar risks [98]. They tend, for example, to underestimate the chances of future personal risks such as getting cancer, alcohol dependence, asthma or AIDS, and overestimate their chances of success in business or winning a lottery [110]. In memory tasks, remembering and retrieval can be so selective that "memory illusions" arise that are similar to perceptual illusions [93]. Salient elements color judgments, issues with emotional or social relevance being more attended to and recalled [79,80].

When asked to make estimates from samples of data, people tend to be insensitive to previous outcomes and probabilities and to sample size. We commonly imagine too small a range of possibilities, get stuck near a starting value rather than range freely, stick to defined values too closely, and create fallacious assumptions based on limited and partial confirmation [87,108]. Many errors in judgment seem to stem from overreliance on heuristics (cognitive shortcuts) that are usually adaptive in problem solving [108]. The representativeness heuristic, for example, evaluates how similar an event is to its precursors and how much it reflects the salient features of the process or the category from which it comes. The availability heuristic relies on the most easily available or retrievable examples. In many cases, when information is incomplete or time is short, reliance on these heuristics is adaptive [6,33]. For example, similarity to other events is often the only reliable criterion for judging the probability of an event. At the same time, reliance on biased heuristics may lead to the ignoring of relevant information and to mistaken conclusions. Such crude and quick judgments seem to satisfy environmental demands but can lead to distorted outcomes.

Experts trained to rule out bias are not immune to such errors, and outside their own area of expertise (and often inside it) they show the same biases as lay people. An example [69] is the widespread one of overconfidence noted earlier: people are overconfident in estimating their performance and the accuracy of their assessments. This bias resists change, persisting despite express warnings against it [59]. Overconfidence could result from a confirmation bias. If people only try to support their favorite hypothesis, disregarding contrary evidence, they will overestimate the chance that their hypothesis is correct [61].

2.3. Experience and judgment biases

Experience may underlie some biases of judgment, as with the "base-rate fallacy" in which we neglect the overall frequency of events when asked to predict their occurrence from the presence of certain cues. In a study [43], subjects had to guess the frequency at which they saw each of two colored patches (green or blue) appearing repeatedly on a

computer screen, after being told at the start the different base-rates of appearance of each colored patch. Subjects' guesses differed according to whether they had seen (i) color-related cues (green or blue *marks*) or (ii) the *words* "green" or "blue" or (iii) white horizontal or vertical lines, preceding the appearance of the target patch. A base-rate error—excessive matching of cues to target—did not appear after the lines, but only after the blue or green marks or the words 'blue' or 'green'. Thus past familiar color or meaning primed appearance of the base-rate fallacy, showed a reasoning preference arising from weighted processing of physically or functionally related features. The rule of associating cues by similarity, proximity, salience or relatedness pervades many types of cognitive processing and underlies reasoning outside awareness as opposed to rule-based conscious reasoning [97].

Base-rate neglect and other cognitive fallacies or illusions appear less when problems are formulated as frequencies rather than as probabilities [21,39]. This may be because people observe frequencies in nature, not probability statements. Nevertheless, to estimate by frequency, too, is a systematic preference that can distort, as has been found with different tasks [59]. Despite Koehler's [63] contention that the base-rate fallacy has been exaggerated, and disputes over its size and relevance, the fallacy still appears across different settings [40,59].

Thus, although judgment biases are usually said to reflect experience (e.g. language rules, thinking habits, contextual factors, education), they tend to appear quickly in typical forms (e.g. jump to distorted conclusions from salient but partial features, exaggerate boundaries and contrasts, overgeneralize from regularities) without explicit training in many sorts of cognitive tasks.

3. Social biases

During social interaction people attend to, perceive, associate and retrieve socially relevant information. Those processes require the operation of neural modules to extract and correlate important features of social scenarios using strategies like those seen in simpler cognitive tasks (attend to salient elements, ignore redundant or monotonous zones, associate along relevant paths, construct "maps" and "search images", etc.) and other strategies that enhance efficiency at the cost of bias. We hypothesise that some social biases depend on properties of those neurocognitive modules which mediate the perceiving and categorizing of socially relevant cues and maps (e.g. voice, gaze, face, gender, size, emotional expression, dominance signals). If the hypothesis is right, the tendency to be biased in social judgments may originate in distortions imposed by neural devices dedicated to target and learn about the social world. Before examining this hypothesis further, some aspects of social biases require summary.

3.1. Social stereotyping

Social categorization and stereotyping have a crucial role in promoting prejudice and social conflict. Social stereotypes form easily. Negative stereotypes of Black Americans formed following a single negative interaction with a black person; moreover, simply overhearing someone say that a black person committed a crime led to group stereotyping and ingroup favoritism [53]. Quickly formed stereotypes influence our perception of, attitudes about and actions towards people.

Facts about people tend to be remembered in ways that fit the stereotype. Recall is better for facts that support rather than disconfirm stereotypes [38], and for negative rather than positive facts, which fits the finding that negative stereotypes outnumber good ones [56]. Recall is actually enhanced by stereotyping, but at the price of losing detail and accuracy [11,100]. Evidence inconsistent with stereotypes is more readily ignored or discarded than supportive evidence [48]. Belief in the validity of stereotypes is propped up by "illusory correlations" [16]: when we expect features to be correlated we perceive covariations which do not exist in the data. False theories are thus reinforced by false observations. These processes work through biased heuristics and contribute to "illusions of validity" giving an erroneous sense of certainty [108]. In brief, facts that contradict rather than confirm our views are attended to less, filtered more strictly, and erased or discarded more easily [14,87]. Thus social stereotypes arise out of and are strengthened by biased cognitive processing.

In studies of categorical perception, humans learn to distinguish between sets of objects and invent illusory properties of those sets. After being trained by trial and error plus feedback to categorize difficult items (e.g. computer-generated textures or chick genitalia), subjects exaggerate within-category similarities and between-category differences [4]. This effect also appeared when identifying identical twins [104] and morphed faces of celebrities [7]. This learning dependent phenomenon of compression/separation in categorical perception may relate to rules for efficient grouping and discrimination of visual [96] or verbal objects [3], and underlies typical exaggerations of stereotyping.

Stereotyping has value in a world where snap judgments about people and groups must often be made before there is time to get to know them well [34]. Indecision due to cautious doubt and waiting for detailed information can end in missing the boat. To save time we "automatically" perceive and learn about others from salient features or pattern—size, shape, skin color, movement, voice, gaze, facial appearance and other aspects that create individual profiles (just as when we detect contours and shapes of visual objects, or rhythms and segments in acoustic objects). These instant maps are melded with trained categories (e.g. cultural labels and schemas) to construct socially primed categories like gender, age, ethnicity, status and profession [14].

3.2. Ingroup biasing

Like stereotypes, ingroup biases appear easily and without explicit training [105]. We quickly form social groups and divide the world into 'friends' and 'foes', those for us and those against us. In a classic 'minimal-group paradigm', the criteria for membership were as arbitrary as a preference for Klee's over Kandinsky's paintings [12]. A single label thus sufficed to promote ingroup biasing. Subjects in another study favored members of a new group who shared their initial preference, immediately rating just-met fellow members of an arbitrary group as being more cordial, fair, honest, reliable and intelligent than nongroup members [95] and more deserving of resources if goods had to be shared [15]. The same applied to children aged 5 years tested in minimal-group paradigms [14,116].

Ingroup biasing has a further outcome. It fosters a sense of social identity that binds the group together [105]. Personal identity expands to include the group. Part of the shared identity may come from a generalization of self-serving individual biases. Part may also come from an automatic accentuation of between-group differences and minimization of within-group ones—an example of the compression/separation of categorical perception [30]. Such preparedness for quick categorization and binding/identification inside groups, with a resulting tendency to share ideas and goods, probably yields benefits from group membership that may promote survival [113,114].

The bias for cooperative behavior [15] also incurs a dangerous cost—prejudice against the outgroup, with a potential for social conflict. There may be no alternative, however. We inevitably perceive social categories relevant to our econiche, think in stereotypes and join groups, at least temporarily, because of the human evolved need for 'group-ishness' [92].

4. Links between social and neurocognitive biases: a working hypothesis

We hypothesize that social stereotypes arise easily and perhaps unavoidably from the automatic operation of neurocognitive modules for social adaptation. Quick categorizing, and swift generalizing across situations, speed decisions that have to take into account features such as age, gender, ethnicity and status. It is equally important to grasp quickly cues that predict cheating, honesty, loyalty and animosity [35,92]. Such social judgments require priming, as they are so complex. They can require, for example, the detection and correlation of current and past facial expressions, gestures, speech, tone, posture and other emotional clues. Let us turn to evidence for the preparedness of social prejudice and its possible neural roots.

4.1. Early appearance of social prejudice

Young children's predisposition to categorize and stereotype by gender, race or age supports our idea that neurocognitive modules promote the appearance of prejudice. 3-year-old white-American children constructed "race-based identities" with little or no prompting [55]. When asked to match pictures of adults and children differing in conspicuous external features, the 3-year-olds matched more quickly using racial similarities than using other conspicuous similarities of occupational dress or body build. The 3 year olds' preference for racial matching did not depend on skin color alone and appeared despite their ability to categorize by occupational dress or body size. Another group of white-American 3 year olds was asked to match the sound of an unfamiliar (Portuguese) or familiar (English) language to drawings of black vs white people, exotic vs typical American clothes, and exotic vs typical American homes. The children matched the sound of Portuguese more to the pictures of black people and exotic clothes and homes, despite the fact that black-Americans speak English.

Other research from different countries and cultures confirmed that children as young as (2–2.5) years prefer categories of ethnicity, gender and age reviewed in Ref. [14]. In studies of habituation–dishabituation to pictures, babies aged 6 months were already aware of sex and age categories [99]. This ability develops almost in synchrony with their capacity to distinguish categories of sound, color, and shape, detection of which involves specialized neurocognitive modules. When looking at pictures of adult faces that were attractive or unattractive (as classified previously by adults), babies as young as 2–3 months consistently attended more to the pictures of attractive faces [65,66]. This suggests that the brain has systems guiding attention to cues associated with attractiveness; these may relate to preferences for symmetry (see below).

There is thus much evidence that an ability to detect socially salient cues and profiles emerges very early in life and that children use them without prior training to do so. This early appearance of social biases suggests that they depend on inborn or early molded neurocognitive modules that guide propensities to categorize along such lines. The involvement of neurocognitive modules does not exclude a role for learning. Such modules often *require* some learning for their expression, without which the ability may not emerge. The point is that the modules expedite *rapid* learning of specific abilities at a given age, just as other modules permit the rapid acquisition of speech in young children provided they hear people speaking around them. The ability of 8-month-old babies to extract iterative segments of "language" from a continuous stream of pseudosyllabic nonsense [102] or of 7 month olds to extract abstract algebraic rules from an artificial language [70], are examples of the operation of such learning expediting modules.

4.2. The propensity to categorize

Current models assume that the brain evolved to construct and assemble workable categories out of an unlabeled,

uncertain and ever changing world [17,29,47]. Mainstream research in social psychology concluded that the appearance and persistence of prejudiced thinking is largely explained by the human tendency to categorize and generalize from particular cues, prompts or labels [14,87]. In other words, prejudice is one manifestation of a basic propensity to categorize.

To categorize is to sort things according to certain features (e.g. color) while neglecting others (e.g. shape). Selecting by some aspects while ignoring others that may be relevant exemplifies prejudice that can lead to mistaken sorting. People categorize by looking for 'family' resemblances (typicalities) in ways that lead to groupings by particular features [29,49,73]. This does not depend specifically on language. Pigeons constructed categories like 'tree' and 'faces' and generalized from them just by repeated observation without instrumental learning [54]. In doing that, the birds discriminated essential from similar but inessential features and ignored great variations in context. Further, when neural nets are required to categorize, they develop the same within-group compression and between-group separation bias that humans do [50].

There are, thus, prerequisites for social concepts to be sought in brain modules that process social stimuli and their operational modes that fuel a tendency to categorize and generalize. Social labels build on categories like gender, age, skin color, size, attractiveness and friendliness, which require efficient processing of perceivable features. Fox [34] conjectured: "as long as perceivable differences exist we can only hope to revise our stereotypes in a more favorable direction, not try to outlaw what is not a disease of the mind but part of its basic constitution. We have to come to terms with the idea that prejudice is not a form of thinking but that thinking is a form of prejudice". We temper this by saying that much, not all, thinking uses inherently biased (prejudiced) ways of reasoning [97]. Fox is right in holding that perceivable differences are the frame whereby the brain captures and organizes an unlabeled world (physical and social) from infancy onwards.

4.3. Specific modules for social cognition?

Faces and their emotional expressions are vital social cues. Specialized brain cells recognize and store faces in the inferotemporal cortex in monkeys [27,82] and in corresponding regions in humans. Neuroimaging studies found modules in the human brain that recognize facial emotional expressions, of fear and anger patterns primarily in the amygdala [1,2] and of disgust in the insula and other regions that mediate the appraisal of offensive tastes [85]. These modules function automatically outside awareness though their behavioral and physiological effects are measurable [75,112]. Research still has to establish for each task, which systems deal specifically with socially relevant cues and which are recruited in synchronous but nonspecific scanning/targeting.

The amygdala seems to mediate judgment of other people's likely social behavior from their facial expressions. This was suggested by a study of three patients with bilateral amygdala damage [2]; in a comparison with normals and with other brain-damaged patients they showed defective recognition of facial emotional expressions and defective judgments of the two social attributes of approachability and trustworthiness. Compared to the other subjects, the three amygdala-damaged patients rated unfamiliar people as being far more approachable and trustworthy; differences were greatest with those faces that the normals rated as the most negative. Such changes in social appraisal did not extend, however, to judgments of people's approachability and trustworthiness based on verbal descriptions. Nor was it due to an ignoring of obvious facial features (direction of gaze, expression of the eyes and mouth, visibility of the eyes).

The above evidence directly supports our claim that particular neural modules help to mediate socially relevant cognitions. That the verbal domain was unaffected may point to specificities in the paths recruited in cognitive processing. This finding seems to go against our assumption of continuities in biasing across different levels of processing, but this issue requires direct study. For instance, after backward masking conditioning of facial emotional expressions outside awareness [75,81], a study could be made in normals of the tendency to generalize or stereotype across levels e.g. photographs vs verbal descriptions.

5. A prerequisite: neural modules mediating cognitive biases

We assumed that common biases in attending, perceiving, associating, retrieving, assessing and other cognitive functions arise from the routine operation of specific neural modules that evolved for efficient appraisal and adaptive action in the world. Most of these routines are inbuilt or forged early in development and they work automatically outside conscious scrutiny. Evidence for lower-level roots in cognitive strategies leading to bias comes from commonalities in operational modes across domains (e.g. stimulus relevance in conditioning, dedicated circuits to segment "objects" in visual stimuli, resolution windows and speed limits in attentional tasks, physical regularities and temporal relatedness in generalizing across events etc.).

Further evidence is the relevance of emotional processing for promoting adaptive judgment. That link is illuminated by a recent neurological finding that covert emotional biases play a role in judgment. This was shown in studies of gambling for real money by normal subjects and by patients who had bilateral lesions in the ventromedial region of the prefrontal cortex [8]. The normal subjects experienced physiological reactions associated with "hunches" or "intuitions" that warned them against high-risk choices and guided them towards making adaptive (winning) choices,

well before they could formulate and explain verbally the rules of the gamble. The conscious reasoning of the normal subjects was thus preceded by nonconscious biases favoring advantageous behavior. In contrast, the patients did not have the physiological reactions and hunches of the normals. Despite becoming able to discover and state the rules of the gamble *at the same time* as the normals, the patients persisted with high-risk choices. Their lesions had apparently damaged the nonconscious neural biases that benefited the normals [8,9,24]. This exemplifies an intuitive loose system of reasoning coming to the aid of conscious rule-based reasoning to guide adaptive options [97]; also see Refs. [33,68] for related ideas.

Damasio suggested that such emotionally charged biases (automatic warnings from prefrontal ventromedial sites that store the approximate punishment/reward ratios obtained in a task) aid the work of other modules whose integrity is crucial for decision-making in humans and other primates. These modules involve response–inhibition systems for delaying answers, short-term working memory [36] and selective attention routines [52,88]. Together, these large and distributed neural systems mediate a variety of processes that are inherently biased and operate outside awareness. However, the formal rules of objective and conscious reasoning also have particular neural substrates with characteristic constraints. There is evidence that logical rules in reasoning such as *if…then* [57], and the processing of number (quantity) and elementary arithmetic [26] depend on specific domain representations linked to the work of particular brain systems.

5.1. Evolving bias in artificial neural networks: examples in social signaling

Bias appears as an adaptive byproduct not only in animals but also in artificial neural networks that are allowed to "evolve" over successive generations. Spontaneous bias appeared in simulations of female birds' preference for exaggerated male traits (bright colors, elaborate ornaments and conspicuous displays) [31]. The network modeled a female visual recognition system, which had to distinguish male conspecifics from males of other species differing only in their tail length. The system was tested in an iterative procedure mimicking how recognition systems may change by natural selection over evolutionary time. Several solutions evolved rapidly to produce near error-free recognition.

Surprisingly, some novel stimuli that had not been used during training but which exaggerated aspects of the conspecific male (overlong or bizarre tails) evoked stronger responses than did the original stimulus; these became supernormal stimuli. The network's spontaneous bias for exaggerated features increased after further evolutionary simulations allowed mutation in male tail length and in the female recognition system. The shift in the sensory responsiveness of females towards long or otherwise conspicuous male tails continued even when they threatened

male adaptiveness, thus confirming Darwin's view that a trait conferring an advantage in mating can evolve to such an extreme that it shortens male survival.

Spontaneous bias also appeared in a net trained to model female preferences for symmetry in male ornamentation [58]. The net was "trained" to distinguish images of bilateral "tails" that exhibited varying degrees of asymmetry. Selection in a repeatedly mutating network led to a clear preference for more symmetrical figures and an adjunctive preference for longer tails too. Thus, results indicated that when a male of a species has paired ornaments, selection for female recognition of appropriate mates could lead to evolved biases favoring symmetrical males. In further work, neural networks evolved to prefer symmetrical images and colors out of random patterns [32]. The networks had to react to colored squares in a grid presented in different positions and orientations so as to mimic the way animals see objects in their visual field. When coevolution of signals and of networks was allowed, systematic preferences appeared for symmetry and also for purer and brighter colors.

Geoffrey Miller (personal communication) was unable to replicate the above results using visual patterns, but obtained preliminary evidence of neural networks evolving to show spontaneous bias in other interactive tasks such as sexual selection of rhythmic acoustic displays. Enquist and Arak [31] suggested that a biased response to signals, arises inevitably since a target signal can take an almost infinite number of forms and a recognition mechanism must always prefer some variants in order to respond efficiently. A bias for relevant and conspicuous features may thus be inherent in any signal detection/response system. Further work on the sexual selection of male displays in several species reinforced this hypothesis to explain the origin of female preferences and the diversity of male signals. Such studies complemented others that trained artificial networks to mimic evolutionary transitions in the recognition "windows" of species-specific mating calls in frogs [94]. The networks ended with the same type of biases as living frogs. Thus, the inherent neurocognitive biases of receivers play a pivotal role in a coevolving pattern of preferences in social communication systems.

In brief, both artificial and living neural networks evolve biases which are adaptive in a variety of situations (not just those for which the systems first evolved). Such inherent biases manifest in related outcomes that can be found across different levels of cognitive processing including social cognition. The appearance of consistent preferences for salient features or frames is crucial for efficient performance in scanning, associating, comparing and decision making. In abstracting signals from an unlabeled world, the brain must select relevant cues (edges, boundaries, contrasts, segments, forms, frequencies, rhythms etc.) out of the ever-changing sensory input that assails it at every moment. Which features will be selected as relevant depends on the interplay between genetic programs and environmental inputs

(especially in early development) with relative weights that need to be established for every task.

6. Avenues for neuroscience research at the edge of social cognition

A challenge for neuroscience is how to identify the brain systems, which mediate the various biases that have been analyzed at the functional level. Neuroimaging methods permit the study of cognitive performance while neural activation is measured. For example, using PET scans [90] attentional networks in the brain mediating object categorization in visual tasks, which required attention to single vs conjoined features were identified. One attentional network was more implicated in top–down processes linked to the detection of conjoined visual features during the recognition of ellipses of different colors or orientations, while another brain network was more related to sustained attention not directly dependent on stimulus changes. Hence, there is room for distinguishing top–down vs bottom–up biases in attentional tasks along with the particular neural templates associated with them.

The way is also open now for the study of neural templates of perceptual and memory illusions [106] and conjunction fallacies [89], to cite just two examples of biased processing in tasks that lie at the boundary of object- and social cognition. Such studies need to explore the neural correlates of categorical perception [49] and of comparison processes leading to similarity judgments [73]. The level of study we propose for a fruitful analysis of the neural substrates of cognitive biasing is that of systems/networks [17,18,36,62]. Cellular, synaptic and molecular levels may be ignored provisionally in accounting for human cognitive biases. Future developments may reveal a need for such downward excursion (e.g. systematic emotional biases probably depend on typical neurohormonal profiles in particular brain regions, and they contribute to biases in judgment).

6.1. Future research to uncover the roots of bias

Several lines of research at the boundary of social psychology and neuroscience could illuminate the role of lower-level neurocognitive modules in generating prejudiced thinking and behavior.

1. Within the same cognitive domain, modifying crucial features or frames at one level would have consequences at another. For example, at the perceptual level exaggerating saliencies, masking contours or colors, confounding similarities, accentuating/reducing symmetry etc should consistently change the outcomes at the level of social cognition and accompanying biases. Changes in facial shape to be used in recognition tasks (perceptual level) should lead to changes in higher levels, e.g. attractiveness [83], dominance/submission and other attributes that affect stereotyping. Similar manipulations could be studied for voice, size etc.
2. Neuroimaging studies should show regional differences in brain activity during the recognition of visual patterns or tonal profiles of social categories (e.g. gender) that form typical frames for stereotyping.
3. Patients with lesions that impair recognition of social cues (e.g. of anger, disgust, symmetry of appearance, accent) should show less tendency to stereotype in the relevant domains (the study of (2) is a first step in that direction).
4. Systematic deviations in lower-level tasks should predict bias in some higher-level cognitive tasks. For example, performance in perceptual-illusion tasks involving figures, shapes etc should predict distortions in the categorizing of related social stimuli. Generalization gradients from visual or tonal stimuli should predict the gradients from related social cues.
5. Within artificial neural networks, lower-level rules should influence the evolution of higher-level biases. If a network is given rules that are basic to human object recognition and categorization (e.g. attend to salient features, detect contours, exaggerate boundaries, fill in discontinuities), when it is required to learn social categories it should evolve stereotypes similar to those found in children—gender, size, color (showing spontaneous computer 'racism'—a non-PC PC?) etc.
6. It should be possible to identify enduring biasing styles (in perceiving, judging or emotion-appraising tasks) that predict stereotyping styles that may differ by personality type and disorder.
7. Learning new categories by explicit labeling will be expedited to the extent that they associate with social labels such as gender, age, size, color, familiarity, resemblance, ethnicity etc.
8. The absence of certain crucial biases might turn out to be maladaptive. Autistic children may lack capacities for processing socially relevant signals. Training normal children to ignore social cues may impair relevant performance in real life e.g. learning to ignore size could lead them to be unaware that a big bully might beat them up.

7. Evolutionary advantages of bias and prejudice

Different tasks demand different types of bias. Preprogrammed preferences and distorted observations represent types of bias, each of whose particular functions and properties should be clearly distinguished. To that end, research is needed on mechanisms subserving each type of bias. We can outline, however, some strategies that may have long been broadly adaptive for hominids and underlie the tendency to think and operate with biases.

Automatic prejudice probably makes sense where a rigid pre-established program is more convenient than flexible learning [101] e.g. startle responses and innate fears. Such

reactions depend on the automatic activation of emotional systems in the brain [68]. It seems nevertheless paradoxical to have evolved to forget persistently inconvenient—though sometimes very relevant—facts, or to attend to too narrow a range of stimuli or to make decisions according to partially erroneous perceptual maps. Would it not be more adaptive to always hone our model of the world detail-by-detail and second-by-second according to what we find? An answer may be that, given the brain's limited hard-wired and plastic capacities, bias is an unavoidable compromise for efficient (not optimal) working of neural systems mediating mental and behavioral tasks. Without such evolved biases, we could not function in the face of challenges posed by the environment. The cost of being completely open-minded, ingenuous and flexible might be biologically prohibitive. "Better safe than sorry" is a common theme behind many types of bias [41,42].

The first broadly adaptive strategy may be the efficient focusing and detection of targets. The swift and flexible processing needed to chase fast-moving prey which is constantly changing its position demands a hard-wired or early molded neural repertoire to track a 'search-image' and regard all else as irrelevant background so as to free capacity to decide on the exact moment to pounce. Similar constraints are probably required in trying to detect elusive edible nutrients, in anticipating or escaping from danger, in recognizing potential mates or in detecting collaborators/defectors.

A second and slightly different approach is required when exploring terrain to look for prey or food whose whereabouts is unknown. In order to get the big picture, minute details must be ignored and a tracking process focussing on active pursuit or scanning of a target is inefficient. Arriving at a feasible plan of some kind then takes precedence over the potentially paralyzing search for the best possible solution.

Whichever strategy is used, categorization is involved. To categorize is to sort, based on some features and not others. The learning nets of the brain can only extract provisional features and the world decides whether those are right or not. There is no alternative. By categorizing, the brain constructs a knowable and workable world. This may be a third advantage of bias. Appraisals of constancy, familiarity and regularity depend on creating relatively stable categories at all levels from object perception to social cognition. Such categories are the templates for efficient action despite their incurring of unavoidable biases.

The fourth advantage concerns the confirmation bias with which we began. Darwin himself hinted that the pursuit of suggestive but tenuous evidence instead of looking for potential disconfirmation at the outset enhances motivation. It can be called 'judicious blindness'. If disconfirmation discourages us into giving up the testing of a hypothesis prematurely then we may never give that hypothesis enough chance to prove its worth. Many an idea that seems crazy at first sight proves to be a winner on closer testing and scrutiny. A related idea was advanced in Lakatos's [67]

observation that the most successful scientific programs have moved forward undeterred by initial encountering of anomalies.

Self-serving biases in thinking can accordingly have a general driving, motivational, function. It can expedite the exploration of productive ideas in the face of older systems that have led either to exhaustive but futile testing of old ground or to sterile skepticism. Ideological bias of a religious, political or scientific kind can galvanize activity into creative (or destructive) channels. Enthusiasts with a clear (though not necessarily realistic) vision can progress more quickly than sceptics. Some self-serving thinking may thus be necessary to generate the enthusiasm to pursue goals.

The hypothesis that biased thinking may enhance motivation is supported indirectly by studies of the cognitive styles of depressed people. Depression involves an impairment of the neural mechanisms of motivation with accompanying (and abnormal) negatively biased attention and thinking [13,111]. In studies of visual attention tasks, depressed people scanned their environment more exhaustively [44], thus agreeing with the notion that a depressed mood may perhaps yield a more realistic view of the world, while normal people use narrow, optimistic and openly self-serving biases. Beyond a certain point, of course, severe depression may lead to paralyzing gloom that is maladaptive. However, some forms of depression improve with psychotherapies, which teach problem-solving [76] or try to change negative to positive thinking [10]. The subtle euphoriant effects of normal biased thinking may thus facilitate adaptive behavior.

8. Ideological traps

Some critics have wrongly dismissed cognitive biases as 'artifacts' because they can be modified by context, training, gender or social culture [63]. This is like dismissing the presence of a disposition for learning a language as an artifact merely because *which* language we learn depends upon the language we hear as we grow up. Progress would be faster if emphasis was instead placed on biases which are relevant and widespread, on the underlying neurocognitive modules they reflect and on the distortions induced.

We anticipate objections to our rather unpalatable notion of a partially inborn or early molded "prejudiced" brain. Testimony to the importance of holding onto our cherished ideas and notions is the pain we often feel when evidence piles up on us to abandon them. It can be quite a struggle to give them up. Why is this so? A theory (an elegant chain of formal assumptions and deductions, an intellectually elaborated stereotype or a plain prejudice) tells us what to do in a variety of situations. If it is wrong then we need better guides, and that may be hard to get. Until we find new ones we are at more risk in acting in an uncertain world. This may explain why it can be so painful to have to reappraise major and minor beliefs. Idols (fashion, rituals,

religions, ideologies and rigid scientific doctrines) have a function. Our deep-rooted neurocognitive biases helps them to thrive.

8.1. Prejudices need not be immutable

Saying that we are prone to develop prejudice for "friends" and against "foes" does not mean that this inclination cannot be mitigated, provided we recognize it, understand how it happens and guard against it becoming extreme. Many reject evidence that humans are xenophobic, as other species are, for fear that it will be seized on by the political right to justify racist practices, just as Social Darwinism and eugenics were misused earlier this century. Equally, liberals seize on evidence that prejudice can be modified to build false hopes that we could educate away all tendency to prejudice. Each side of the political spectrum sees what it wants to see and rejects an alternative possibility—careful delineation of how intolerance originates from our phylogenetic heritage of specific neurocognitive modules interacting with environmental input from the womb onwards.

Uncovering the roots of persistent biases in judgment and social life has already begun with recent progress in integrating cognitive neuroscience, behavior genetics, human ecology and evolutionary biology [115]. We concur with Wilson that "human nature is not the genes which prescribe it or the universals of culture which are its products. It is rather the epigenetic rules of cognition, the inherited regularities of cognitive development that predispose individuals to perceive reality in certain ways and to create and learn some cultural variants in preference to competing variants. Epigenetic rules have been documented in a diversity of cultural categories, from syntax acquisition, color vocabularies, cheater detection and others. The continuing quest for such inborn biasing effects promises to be the most effective means to understand gene-culture coevolution and hence to link biology and the social sciences causally".

The task ahead is to deepen this knowledge. A parallel task is to find ways to engage in the political debate without being silenced. Such engagement could empower educational and other social institutions to try to reduce enduring problems caused by the prepared tendency of humans, like other species, to behave with self-serving (and its extension, group-serving) biases. A detailed map is needed of when, what and how dangerous social biases emerge, and how each one might be reduced. Such lessening of biases will require effort, but think how much effort and specialised tools are spent on teaching unnatural skills like reading and writing, with gratifying results in the long run.

Acknowledgements

References

[1] Adolphs R, Tranel D, Damasio H, Damasio A. Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. Nature 1994;372:669–72.

[2] Adolphs R, Tranel D, Damasio A. The human amigdala in social judgement. Nature 1998;393:470–4.

[3] Altmann GTM. The ascent of Babel: an exploration of language, mind and understanding, Oxford: Oxford University Press, 1997.

[4] Andrews J, Livingston K, Harnard S, Fischer U. Are concepts grounded in categorical perception? Some relevant empirical results. Paper submitted to the Annual Meeting, Society for Philosophy and Psychology. Memphis, TN; 1994.

[5] Baddeley A. Working memory. Science 1992;255:556–9.

[6] Baron J. Thinking and deciding, Cambridge, MA: Cambridge University Press, 1994.

[7] Beale JM, Keil FC. Categorical perception as an acquired phenomenon: what are the implications? In: Smith L, Hancock P, editors. Springer workshops on computing series, Berlin: Springer, 1996.

[8] Bechara A, Damasio H, Tranel D, Damasio AR. Deciding advantageously before knowing the advantageous strategy. Science 1997;275:1293–5.

[9] Bechara A, Damasio H, Tranel D, Anderson SW. Dissociation of working memory from decision making within the human prefrontal cortex. J Neurosci 1998;18(1):428–37.

[10] Beck AT, Rush AJ, Shaw BF, Emery G. Cognitive therapy of depression, New York: Guilford Press, 1979.

[11] Bellezza FS, Bower GH. Person stereotypes and memory for people. J Personality Social Psychol 1981;41:856–65.

[12] Billig M, Tajfel H. Social categorization and similarly in innergroup behavior. Eur J Social Psychol 1973;3:27–52.

[13] Bradley BP, Matthews A. Memory bias in recovered clinical depressives. Cognition Emotion 1988;2(3):235–45.

[14] Brown RJ. Prejudice, Oxford: Blackwell, 1995.

[15] Coporael LR, Dawes RM, Van de Kragt Orbell AJC. Selfishness examined: cooperation in absence of egoistic incentives. Behavioral Brain Sci 1989;12:683–739.

[16] Chapman LJ, Chapman JP. Illusory correlations as an obstacle to the use of valid psychodiagnostic signs. J Abnorm Psychol 1969;74:271–80.

[17] Churchland PM. The engine of reason, the seat of soul: a philosophical journey into the brain, Cambridge, MA: MIT Press, 1993.

[18] Churchland PS, Sejnowski TJ. The computational brain, Cambridge, MA: MIT Press, 1993.

[19] Clark RE, Squire LR. Classical conditioning and brain systems; the role of awareness. Science 1998;280:77–81.

[20] Cosmides L, Tooby J. From function to structure: role of evolutionary biology and computational theories. In: Gazzaniga MS, editor. Cognitive neurosciences, Cambridge, MA: MIT Press, 1995. p. 1199–210.

[21] Cosmides L, Tooby J. Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgement under uncertainty. Cognition 1996;58:1–73.

[22] Crick F. The astonishing hypothesis, New York: Scribner's, 1994.

[23] Crick F, Koch C. Are we aware of neural activity in primary visual cortex? Nature 1995;375:121–3.

[24] Damasio AR. Descartes' error: emotion, reason and the human brain, New York: Grosset/Putnam, 1994.

[25] Darwin C. In: Darwin F, editor. The autobiography of Charles Darwin, 1876, New York: Appleton, 1892 Reprinted 1958, Dover, NY.

[26] Dehaene S, Dehaene-Lambert G, Cohen L. Abstract representations

of numbers in the animal and human brain. Trends Neurosci 1998;21(8):355–61.

[27] Desimone R. Face-selective cells in the temporal cortex of monkeys. J Cognitive Neurosci 1991;3:1–24.

[28] Driver J. Enhancement of selective listening by illusory mislocation of speech sounds due to lip-reading. Nature 1996;381:66–8.

[29] Edelman GM. Neural Darwinism, New York: Basic Books, 1987.

[30] Eiser JR. Accentuation revisited. In: Robinson P, editor. Social groups and identities: developing the legacy of Henri Tajfel, Oxford: Butterworth Heinemann, 1996.

[31] Enquist M, Arak A. Selection of exaggerated male traits by female aesthetic senses. Nature 1993;361:446–8.

[32] Enquist M, Arak A. Symmetry, beauty and evolution. Nature 1994;372:169–72.

[33] Epstein S. Integration of the cognitive and the psychodynamic unconscious. Am Psychol 1994;49(8):709–24.

[34] Fox R. Prejudice and the unfinished mind: a new look at an old falling. Psychol Inquiry 1992;3:137–52.

[35] Frank RH. Passions within reason: the strategic role of emotions, London: Norton, 1998.

[36] Fuster JM. Memory in the cerebral cortex, Cambridge, MA: MIT Press, 1995.

[37] Garcia J, Lasiter PS, Bermúdez-Rattoni F, Deems DA. A general theory of aversion learning. In: Braveman NS, Bronstein P, editors. Experimental assessments and clinical applications of conditioned food aversions, *Ann NY Acad Sci*, 443, 1985. p. 8–21.

[38] Gergen KJ, Gergen MM. Social psychology, New York: Springer, 1985.

[39] Gigerenzer G. In: Stroebe W, Hewstone M, editors. How to make cognitive illusions disappear: beyond heuristics and biases, *Eur Rev Social Psychol*, 2. Chichester: Wiley, 1991. p. 83–115.

[40] Gigerenzer G. On narrow norms and vague heuristics: a reply to Kahneman and Tversky. Psycholo Bull 1996;103(3):592–6.

[41] Gilbert P. Evolutionary psychopathology: why isn't the mind designed better than it is? Br J Med Psychol 1998;71:353–73.

[42] Gilbert P. The evolved basis and adaptive functions of cognitive distortions. Br J Med Psychol 1998;71:447–63.

[43] Goodie AS, Frantino E. Learning to commit or avoid the base-rate error. Nature 1996;380:247–9.

[44] Gotlib IH, McLachlan AL, Katz AN. Biases in visual attention in depressed and nondepressed individuals. Cognition Emotion 1988;2(3):185–200.

[45] Gregory RL. Eye and brain: the psychology of seeing, 3. New York: McGraw-Hill, 1978.

[46] Gregory RL, editor. Oxford companion to the mind New York: Oxford University Press, 1987.

[47] Grossberg S, Mingolla E, Ross WD. Visual brain and visual perception: how does the cortex do the perceptual grouping?. Trends Neurosci 1997;20(3):106–11.

[48] Hamilton D, Dugan P, Troiler T. The formation of stereotypic beliefs: further evidence for distinctiveness-based illusory correlations. J Personality Social Psychol 1985;48:5–18.

[49] Harnad S, editor. Categorical perception: the groundwork of cognition New York: Cambridge University Press, 1987.

[50] Harnad S, Hanson HJ, Lubin J. Learned categorical perception in neural nets: implications for symbol grounding. In: Honavar V, Uhr L, editors. Symbol processors and connectionist network models in artificial intelligence and cognitive modelling: steps toward principled integration, New York: Academic Press, 1995. p. 191–206.

[51] He S, Cavanagh P, Intriligator J. Attentional resolution and the locus of visual awareness. Nature 1996;383:334–7.

[52] He S, Cavanagh P, Intriligator J. Attentional resolution. Trends Cognitive Sci 1997;1(3):115–21.

[53] Henderson-king EI, Nisbett RE. Antiblack prejudice as a function of exposure to the negative behavior of a single Black person. J Personality Social Psychol 1996;71:654–64.

[54] Herrnstein RJ. Stimuli and the texture of experience. Neurosci Biobehavioral Rev 1982;6:105–17.

[55] Hirschfeld LA. Race in the making. Cognition, culture and the child's construction of human kinds, Cambridge, MA: MIT Press, 1996.

[56] Howard JW, Rothbart M. Social categorization and memory for ingroup and outgroup behavior. J Personal Soc Psychol 1980;38:301–10.

[57] Johnson Laird PN. Mental models, deductive reasoning and the brain. In: Gazzaniga MS, editor. The cognitive neurosciences, Cambridge, MA: MIT Press, 1995. p. 999–1008.

[58] Johnstone RA. Female preferences for symmetrical males as a by-product of selection for mate recognition. Nature 1994;372:172–5.

[59] Kahneman D, Tversky A. On the reality of cognitive illusions. Psychol Bull 1996;103(3):582–91.

[60] Klayman J, Ha YW. Confirmation, disconfirmation, and information in hypothesis-testing. Psychol Rev 1987;94:211–28.

[61] Koriat A, Lichtenstein S, Fischoff B. Reasons for confidence. J Exper Psychol: Human Learning Memory 1980;6:107–18.

[62] Koch C, Davis JL, editors. Large-scale neural theories of the brain Cambridge, MA: MIT Press, 1994.

[63] Koehler JJ. The base-rate fallacy reconsidered: normative, descriptive and methodological challenges . Behav Brain Sci 1996;19(1):1–53.

[64] Kuhl PK, Andruski JE, Chistivih IA, Chistovich LA, Kozhevnikova EV, Riskina VL, Stolyiariva EI, Sundberg U, Lacerda F. Cross-language analysis of phonetic units in language addressed to infants. Science 1997;277(5326):684–6.

[65] Langlois JH, Roggman IA, Casey RJ, Ritter JM, Riesen-Donner IA, Jenkins VI. Infants preferences for attractive fades; rudiments of a stereotype. Developmental Psychol 1987;23:363–9.

[66] Langlois JH, Ritter JM, Roggman IAM, Vaugh LS. Facial diversity and infant preference for attractive faces. Developmental Psychol 1991;27:79–84.

[67] Lakatos I. Falsification and the methodology of scientific research programmes. In: Lakatos I, Musgrave A, editors. Criticism and the growth of knowledge, Cambridge: Cambridge University Press, 1970. pp. 91–196.

[68] Le Doux J. The emotional brain, New York: Simon and Shuster, 1997.

[69] Lichtenstein S, Fischoff B, Phillips LD. Calibration of probabilities: the state of the art. In: Jungermann H, Dezeeuw G, editors. Decision making and change in human affairs, Amsterdam: Reidel, 1977.

[70] Marcus GF, Vijayan S, Bandi Rao S, Vishton PM. Rule-learning by seven month-old infants. Science 1999;283:77–80.

[71] Marks IM, Tobena A. Learning and unlearning fear: a clinical and evolutionary perspective. Neurosci Biobehav Rev. 1990;14:365–84.

[72] Marr D. Vision, San Francisco: Freeman, 1982.

[73] Medin DL, Goldstone RL, Gentner D. Respects for similarity. Psychol Rev 1993;100(2):254–70.

[74] Mele AR. Real self-deception. Behav Brain Sci 1997;20:91–136.

[75] Morris JS, Ohman A, Dolan RJ. Conscious and unconscious emotional learning in the human amygdala. Nature 1998;393:467–70.

[76] Mynors-Wallis LM, Gath DH, Lloyd-Thomas AR, Tomlinson D. Randomised controlled trial comparing problem solving treatment with amitriptyline and placebo for major depression in primary care. Br Med J 1995;310:441–5.

[77] Näätänen R, Lehtokoski A, Lennes M, Cheour M, Huotilainen M, Ilvonen A, Vainio M, Aiku P, Ilmoniemi RJ, Luuk A, Allik I, Sinkönen J, Alho K. Language-specific phoneme representations revealed by electric and magnetic brain responses. Nature 1997;385:432–4.

[78] Neuhoff JG. Perceptual bias for rising tones. Nature 1998;395:123–4.

[79] Nisbett RE, Ross L. Human inference: strategies and shortcomings of social judgement, New York: Prentice-Hall, 1980.

[80] Nisbett RE, Borgia E, Crandall R, Reed H. Popular induction:

information is not necessarily informative. In: Kahneman D, Slovic P, Tversky A, editors. Judgement under uncertainty: heuristic and biases, New York: Cambridge University Press, 1982. p. 101–16.

[81] Ohman A. Fear and anxiety as emotional phenomena: clinical phenomology, evolutionary perspectives and information processing mechanisms. In: Lewis M, Havilan JM, editors. Handbook of emotions, New York: Guilford, 1992. p. 511–36.

[82] Perret DI, Mistlin AJ, Chitty AJ. Visual neurones responsive to faces. Trends Neurosci 1987;10:38–364.

[83] Perret DI, Lee KJ, Penton-Voak I, Rowland D, Yoshikawa S, Burt DM, Henzil SP, Castles DL, Akamatsu S. Effects of sexual dimorphism on facial attractiveness. Nature 1998;394:335–7.

[84] Peterhans E, Von der Heydt R. Subjective contours: bridging the gap between psychophysics and physiology. Trends Neurosci 1991;14(3):112–9.

[85] Phillips ML, Young AW, Senior C, Brammer M, Andrew C, Calder AJ, Bullmore ET, Perret DI, Rowland D, Williams SCR, Gray JA, David AS. A specific neural substrate for perceiving facial expressions of disgust. Nature 1997;389:495–8.

[86] Pinker S. The language instinct, New York: Morrow, 1993.

[87] Plous S. The psychology of judgement and decision making, New York: McGraw Hill, 1993.

[88] Posner ML, Dehaene S. Attentional networks. Trends Neurosci 1994;17(2):75–9.

[89] Posner ML, Raichle ME. Images of mind, New York: Freeman, 1994.

[90] Rees G, Frackowiak, Frith CH. To modulatory effects of attention that mediates object categorization in human cortex. Science 1997;275:835–8.

[91] Rescorla RA. Pavlovian conditioning: it's not what you think it is. Am Psychol 1998;43:151–60.

[92] Ridley M. The origins of virtue, London: Viking, 1996.

[93] Roediger HL. Memory illusions. J Memory Language 1996; 35(2):76–100.

[94] Ryan MJ. Sexual selection, receiver biases and the evolution of sex differences . Science 1998;281:1299–2003.

[95] Schaller M. In-group favoritism and statistical reasoning in social inference: implications for formation and maintenance of group stereotypes. J Personal Soc Psychol 1992;63(1):61–74.

[96] Singh M, Hoffman DD. Constructing and representing visual objects. Trends Cognitive Sci 1996;1(3):98–102.

[97] Sloman SA. The empirical case for two systems of reasoning. Psychol Bull 1997;119(1):3–22.

[98] Slovic P. Perceptions of risk. Science 1987;236:285.

[99] Small MY. Child development, New York: Harcourt, Brace and Jovanovich, 1990.

[100] Snyder ML, Uranowitz SW. Reconstructing the past: some cognitive consequences of person perception. J Pers Soc. Psychol 1978;36:941–50.

[101] Sober E. The adaptive advantage of learning and a priori prejudice. Paper submitted to HBES Conference. Binghampton, NY, August 1993.

[102] Soffran JR, Aslin RN, Newport EL. Statistical learning by 8 months-old infants. Science 1996;274:1426–8.

[103] Spillman L, Werner JS. Long-range interactions in visual perception. Trends Neurosci 1996;19(10):428–34.

[104] Stevenage S. Which twin are you? A demonstration of induced categorical perception of identical twin faces. Br J. Psychol 1998;89(1):39–58.

[105] Tajfel H. Social identity and intergroup relations, New York: Cambridge University Press, 1982.

[106] Tootell RBH, Reppas JB, Dale AM, Look RB, Sereno ML, Malach R, Brady THJ, Roasen BR. Visual motion-after effect in human cortical area MT revealed by functional magnetic resonance imaging. Nature 1995;375:133–41.

[107] Tononi G, Edelman GM. Consciousness and complexity. Science 1998;282:1846–51.

[108] Tversky A, Kahneman D. Judgement under uncertainty: heuristic and biases. In: Kahneman D, Slovic P, Tversky A, editors. Judgement under uncertainty: heuristic and biases, New York: Cambridge University Press, 1982. p. 3–20.

[109] Van der Heydt R. Form analysis in the visual cortex. In: Gazzaniga MS, editor. The cognitive neurosciences, Cambridge, MA: MIT Press, 1995. p. 365–82.

[110] Weinstein ND. Optimistic biases about personal risks. Science 1989;246:1232–3.

[111] Wells A, Matthews G. Attention and emotion: a clinical perspective, Hove: Lawrence Erlbaum, 1994.

[112] Whalen PJ, Rauch SL, Etcoff NL, McInerney SC, Lee MB, Jenikee M. Masked presentations of emotional face expressions modulate amygdala activity without explicit knowledge. J Neurosci 1998;18(1):411–8.

[113] Wilson DS, Sober E. Reintroducing group selection to the human behavioral sciences. Behavioral Brain Sci 1994;17:585–684.

[114] Wilson DS. Human groups as units of selection. Science 1997;276:1816–7.

[115] Wilson EO. Integrated science and the coming century of environment. Science 1998;279:00 27 March (Essays on science and society series).

[116] Yee MD, Brown RJ. Self evaluation and intergroup attitudes in children aged three to nine. Child Development 1999;63:619–29.