Contents

Arthur S. Reber	102	Some Perhaps Surprising Consequences of the Cognitive "Revolution"
Manfred Hassebrauck	116	The Effect of Fertility Risk on Relationship Scrutiny
Steve Stewart-Williams	123	Darwin and Descartes' Demon: On the Possible Evolutionary Origin of Belief in an External World
Sanjida O'Connell/R. I. M. Dunbar	131	A Test for Comprehension of False Belief in Chimpanzees
Fredric J. Heeren	141	Was the First Craniate on the Road to Cognition? A Modern Craniate's Perspective
Ingo Brigandt	157	Gestalt Experiments and Inductive Observations: Konrad Lorenz's early Epistemological Writings and the Methods of Classical Ethology
Theresa S. S. Schilhab	171	Vertical and Horizontal Learning: Some Characteristics of Implicit and Explicit Learning
Michel A. Hofman	178	Of Brains and Minds: A Neurobiological Treatise on the Nature of Intelligence
Steven M. Platek	189	An Evolutionary Model of the Effects of Human Paternal Resemblance on Paternal Investment
	198	Zusammenfassungen der Artikel in deutscher Sprache

Impressum

Evolution and Cognition: ISSN: 0938-2623 **Published by:** Konrad Lorenz Institut für Evolutions- und Kognitionsforschung, Adolf-Lorenz-Gasse 2, A-3422 Altenberg/Donau. Tel.: 0043-2242-32390; Fax: 0043-2242-323904; e-mail: sec@kla.univie.ac.at; World Wide Web: http://www.kli.ac.at/ **Editors:** Rupert Riedl, Manfred Wimmer **Layout:** Alexander Riegler **Aim and Scope:** "Evolution and Cognition" is an interdisciplinary

forum devoted to all aspects of research on cognition in animals and humans. The major emphasis of the journal is on evolutionary approaches to cognition, reflecting the fact that the cognitive capacities of organisms result from biological evolution. Empirical and theoretical work from both fields, evolutionary and cognitive science, is accepted, but particular attention is paid to interdisciplinary perspectives on the mutual relationship between evolutionary and cognitive processes. Submissions dealing with the significance of cognitive research for the theories of biological and sociocultural evolution are also welcome. "Evolution and Cognition" publishes both original papers and review articles. **Period of Publication**: Semi-annual. **Price**: Annuals subscription rate (2 issues, postage and handling included): Euro 40, US\$ 50, SFr 65, GBP 28. Annual subscriptions are assumed to be continued automatically unless subscription orders are cancelled by written information. **Single issue price** (postage and handling included): Euro 25, US\$ 30, SFr 35, GBP 18. **Publishing House**: WUV-Universitätsverlag/Vienna University Press, Berggasse 5, A-1090 Wien, Tel.: 0043/1/3105356-0, Fax: 0043/1/3197050. **Bank**: Erste österreichische Spar-Casse, Acct.No. 073-08191 (Bank Code 20111). **Advertising**: Vienna University Press, Berggasse 5, A-1090 Wien. **Supported by Cultural Office of the City of Vienna, Austrian Federal Ministry of Education, Science, and Culture, and the Section Culture and Science of the Lower Austrian State Government.**

Some Perhaps Surprising Consequences of the Cognitive "Revolution"

Introductory Remarks

This paper is an attempt to gain some perspective on the impact of the dramatic shift in psychological research and theory that has come to be known as "the cognitive revolution" (BAARS 1986). This intellectual adjustment, which took place over a surprisingly short period of time during the previous century, was marked by many changes in emphasis but none more dramatic that the virtual abandonment of behaviorism for a more subjective, cognitive approach. Most overviews of this 'paradigm shift' have

Abstract

This paper provides an overview of the history of the cognitive revolution that began in the middle decades of the previous century and continues today with unflagging enthusiasm. The author reviews the various moves that accompanied the shift from a scientific psychology dominated by behaviorist thinking to one that takes into account subjective states of mind and emphasizes cognitive functions and their underlying neurological structures. The author, however, argues that this move toward cognitivism was marked by a neglect of two key aspects of behaviorist thought that were passed over by the revolution, specifically the study of learning and the recognition of the importance of DARWINIAN adaptationist principles. Present work seems to be redressing some but not all of the consequences of this dual neglect.

Key words

Cognitive revolution, learning, evolution, adaptationism, memory.

characterized it in distinctly sanguine terms promoting the major advances that emerged as the discipline moved from being a science bound up in a quirky objectivity to one that acknowledged that its most beguiling target was, indeed, an organism marked by subjective, mental experience. And, while sanguinity is appropriate in that the progress made in the past several decades is quite unmatched in the long history of our field, there were two unhappy consequences of the abandonment of behaviorist ways that have gone largely unnoticed. In particular:

(1) The topic of learning fell into disfavor as researchers became more concerned with the mechanisms of memorial representation of knowledge than with the processes that underlie its acquisition. As will become clear, while interest in the topic of *learning* lumbered along in a mundane way with the numposed system have an adaptive role? Could it have evolved in the appropriate time frame?

Neither of these shifts in emphasis was simple and the factors underlying them are still not well understood. What follows is a somewhat personal historical gloss on the decades during which behaviorism lost its purchase in psychological thought and was replaced by a broad-based cognitive science.

The Cognitive Revolution

The much ballyhooed cognitive revolution began, according to most observers, in the 1950s, gathered steam during the 1960s and had pretty much commandeered psychological thought and theory by the middle of the 1980s. BAARS'S (1986) overview of this period, while missing perhaps some of the nu-

Evolution and Cognition | 102 | 2003, Vol. 9, No. 2

ber of studies devoted to this topic reflecting the number of psychologists more than anything else, the study of *memory* and *representation* became increasingly the focus of experimental and theoretical psychology.

(2) The DARWINIAN adaptationist stance that had stood for the better part of a century as psychology's secular talisman lost its influence and ceased to be a part of the psychologist's intellectual quiver. This neglect was particularly serious in that cognitivists began to routinely entertain models without considering the classic DARWINIAN heuristics: Does the proances that were to become important in the following years,¹ nevertheless captured this three-decade shift in thinking quite well. His book, in which historical assessment is engagingly interspersed with short biographical sketches and interviews with significant figures, reflects this sequence of shifts in priorities.

For those of us who began our initiation into experimental psychology during the early phases of this movement, it was a most remarkable time. The dominant behaviorist perspective was under fullbore assault and something, although at the time none of us knew quite what to call it, was beginning to emerge. Behaviorists had tried mightily over the previous four decades to build a psychology that was operational, objective, and free from the kind of "contamination" they felt was brought to science when subjective, unobservable concepts were introduced. Positivism was the touchstone; phenomenology the enemy.

Of course, reality is never so clean as we often would like it to be. And so it was with this "revolution." Cracks in the behaviorists' edifice were actually evident early on. Social instinct theorists such William MCDOUGALL had been launching attacks at radical behaviorism almost from its inception (MC-DOUGALL 1932; WATSON/MCDOUGALL 1928); KOFFKA (1935), KÖHLER (1925) and other Gestalt psychologists had been chipping away at behaviorist thought since their arrival in the United States; Magda ARNOLD (1945, 1970) had been arguing, persuasively in the minds of many, that the causal links that led from emotions to their behavioral consequences only made sense when "appraisal," a most distinctly cognitive element, was incorporated; and even loyalists who identified with the behaviorist's paradigm such as Edward Chace TOLMAN, found themselves introducing quasi-heretical notions like "purpose," "fulfillment," and "cognitive maps" into the mix (see especially TOLMAN 1926, 1932, 1948).

But, as virtually all historians acknowledge, behaviorist thought fought its way past these objections and from roughly the middle 1920s to the late 1950s and early 1960s came to dominate North American experimental psychology in a way that was quite extraordinary. One consequence of this shift was that a broad array of issues that smacked of cognitivist thinking or were deemed by the operative orthodoxy as too dependent on subjective or introspective analysis were no longer regarded as legitimate topics of investigation. It's worth taking a quick look at one of these topics since the forces that made it an early casualty continued to play a role in the manner in which behaviorism staked out and defended its intellectual turf. So:

A Short "Aside" on the Subject of Play

Before the turn of the last century play was a topic of considerable interest (GROOS 1898, 1899) only to largely disappear as behaviorism gained influence. Indeed, except for the occasional critique of the early work (BEACH 1945; SCHLOSBERG 1947), the subject of play virtually disappeared from psychological discourse until the 1970s. FAGAN (1981), in his extensive and fascinating overview of the topic, referred to it as the "ugly duckling of behavioral science" during this period (p33).

While the early work on the topic was, admittedly, seriously lacking in scientific methodology, these scientific shortcomings were almost certainly not the reasons for the decline in interest. It is worth noting that other topics that attracted similar early interest and suffered from similar methodological flaws managed to survive quite well-primarily because their subject matter fit comfortably within the boundaries of behaviorism. For example, the specific topic of learning and the general area of comparative psychology, both of which had been as badly abused by early enthusiasts as play had been (ROMANES 1882/ 1965) and just as vigorously critiqued (MORGAN 1894), managed not only to survive but to become central focuses of behaviorist-inspired research. The key was not a sanitized methodology but the degree to which the subject matter fell within the positivist inspired behaviorist movement.

Play, as a topic of scientific investigation, could simply not be studied without taking into account the mental, emotional states of the participants in the action. And, worse, as FAGAN makes clear, understanding play also involves the supposition that each participant has a representation, not only of the activities involved in the interaction, but of the mental and emotional states of other conspecifics involved in the game. Epistemic extensions of this order were simply not tolerated by behaviorists. Play, unlike learning, could not continue to exist as a topic of investigation under the behaviorist banner and its re-emergence had to await the resurgence of cognitivism. A host of other areas and topics suffered similar fates including major approaches like developmental and child psychology, social and political psychology and health psychology along with more localized topics that were once actively studied such as aesthetics, religion and forensic psychology.²

Evolution and Cognition | 103 | 2003, Vol. 9, No. 2

The 1950s—The Assault on Behaviorism

By the end of the 1940s serious cracks in the behaviorist edifice were becoming all too obvious and these, unlike the gentle murmurings of TOLMAN and ARNOLD, could neither be ignored nor hastily incorporated into behaviorist thought. The criticisms were broad-based, systematic attacks on the very foundations of the paradigm.

The following is a list of candidates for the significant intellectual events that seriously undermined behaviorist dogma. The list is not meant to be exhaustive, merely representative, and, of course, it reflects my particular slant on this period. Some of the contributions were critical in nature, dealing lethal blows to radical behaviorism. Others were more forward looking, inviting scientists to begin to explore new ways of doing psychology. However, as the intellectual movement that stimulated these events unfolded, one can sense, almost palpably, the gradual drift away from concern with issues of learning and the acquisition of knowledge and the virtually complete absence of adaptationist, DARWINIAN considerations.

1. In 1951 Karl LASHLEY's famous article "Serial order in behavior" appeared. This paper was an elegant, subtle and utterly devastating attack on one of the more influential branches of behaviorism, the one championed by Clark HULL of Yale University (HULL 1943, 1951). LASHLEY showed, with a series of clever and compelling examples, that any model based solely on linear mechanisms (as HULL's was) couldn't handle the most obvious of human actions, like speaking a sentence, writing a poem, or playing the piano. HULL, for all intents and purposes, was dead in the water. LASHLEY, ironically, made an interesting intellectual assassin, trained as he was by one John Broadus WATSON.

2. The early 1950s saw the introduction of a number of theories that implied that formalism was not just a faint hope for the future. While HULL had attempted valiantly to use his hypothetico-deductive method to develop mathematically rigorous models (HULL 1952), his approach ultimately collapsed under its own axiomatic weight. The emergence of more sophisticated theories using computer simulation techniques (Newell/Shaw/Simon 1958, Newell/ Simon 1956) showed how neatly mathematically and computationally based systems could mimic human decision making and problem solving behavior. This work had a dramatic and lasting impact. It was very much a part of the information processing approach that was to become the dominant paradigm of the cognitive revolution and, of course, helped pave the way for the emergence of the field of artificial intelligence.

3. Claude SHANNON (1948, 1951; SHANNON/ WEAVER 1949), seeing the deep link between the physicist's notion of entropy and the engineer's concept of information, introduced the notion of information processing to communication theory. Psychologists such as George MILLER (1951, 1956), quickly saw the psychological relevance of information processing and began developing cognitive and linguistic models using these new methods. While many, including MILLER (2003), became wary of overemphasizing this approach, it was quickly picked up by a sufficient number of psychologists such that standard introductory textbooks (e.g., KASSIN 2003) now refer to this era as one dominated by the information processing model.

4. At roughly the same time, a variety of novel mathematical techniques were applied to the problems of learning (BUSH/MOSTELLER 1951, 1955; ESTES 1959), psychophysics (TORGERSON 1958), signal detection (SWETS/TANNER/BIRDSALL 1961) and decision making (LUCE 1959). While these contributions didn't have nearly the direct impact on cognitive theory that the simulation work of Newell, Simon and their colleagues had on the information processing approach, they played an important role in undercutting the impact of radical behaviorism and diverted those who held an interest in formalism away from HULL's approach.

5. In 1954 the Social Science Research Council sponsored an interdisciplinary conference on psychology and language which was attended by an array of significant thinkers from linguistics, anthropology, psychology and related disciplines (see OSGOOD/SEBEOK 1965). The importance of this meeting with its assemblage of young, creative thinkers has not really been appreciated by most historians. Its impact was significant and went a long way toward softening the positivist account of language which had been promoted by prominent theorists like BLOOMFIELD (1933) and had long dominated linguistic work. This conference was followed two years later by a symposium at MIT organized by the then rather revolutionary "Special Interest Group in Information Theory." This meeting had a more limited vision than the SSRC meeting but was equally significant. Interestingly, it was here that CHOMSKY first publicly articulated his novel ideas about transformational grammars. In fact, it has been suggested that this is where the cognitive revolution actually began (MILLER 2003).

Evolution and Cognition | 104 | 2003, Vol. 9, No. 2

6. Jerome BRUNER and colleagues published their seminal *A study of thinking* (BRUNER/GOODNOW/AUS-TIN 1956) outlining a set of heuristic devices that people used in problem solving. This book championed the employment of what became known some decades later as "top-down" cognitive procedures for tackling complex, logical problems.

7. When the decade of the 1950s ended, it did so, not with a whimper, but with a distinct bang. In 1959 a then-obscure linguist named Avram Noam CHOMSKY wrote a rather astonishing demolition of B. F. SKINNER's behaviorist treatise, *Verbal Behavior* (SKINNER 1957). CHOMSKY's critique (CHOMSKY 1959), one of the most influential book reviews in the history of the academy, was a scholarly "dusting up" which quite thoroughly put to rest any lingering hopes for a non-mentalist approach to the problems of language and thought.³

8. The turn of the decade saw the publication of MILLER, GALANTER and PRIBRAM'S (1960) *Plans and the structure of behavior* which outlined how important the establishment of cognitive planning was for virtually everything interesting that people did. This book presented the logical extension of the arguments put forward a decade earlier by LASHLEY (1951) in that MILLER, et al. showed how serial behaviors like speaking, typing, playing the piano, or riding a bicycle, which befuddled behaviorist analysis, could be handled simply by introducing the notion of a *plan*.

9. During the early 1960s psychology discovered, rather belatedly, first Jean PIAGET and then Lev VY-GOTSKY and the entire enterprise of developmental psychology changed overnight. Piaget, a scholar whose work emphasizing the stages of cognitive development had been quite influential in Europe, finally found receptive ears over here across the Atlantic. Vygotsky, whose impact had been largely restricted to the Soviet Union and Eastern Europe, emphasized the role of meaning in language and the importance of social components. Children quickly went from being viewed as organisms characterized by stimulus–response mechanisms to epistemologically sophisticated beings with rich, evolving, cognitive and social systems.

10. North Americans discovered the work of a number of British scholars who, owing to their reluctance to embrace radical behaviorism, had been largely overlooked. Particularly influential were the perspectives of BARTLETT (1932), CHERRY (1953, 1957) and BROADBENT (1958), all of whom encouraged the development of more complex models of perceptual and cognitive functioning.

One of the intriguing features of this shift toward a cognitivist world was the introduction into psychology of the work of scholars trained elsewhere, in other traditions with other values and holding, of course, other prejudices. Psychology was inundated with the ideas and priorities of linguists (CHOMSKY), anthropologists (SEBEOK), mathematicians (LUCE), physicists (BUSH), engineers (SHANNON), communications specialists (CHERRY), computer scientists (NEWELL), statisticians (MOSTELLER), biologists (PIAGET), and political scientists (SIMON⁴). It was all very KUHNIAN and it all happened very quickly. By 1967 Ulric NEISSER was able to title his seminal volume, *Cognitive Psychology* and nobody even blinked.

From a personal perspective, it was an absolute delight to be among the young students of the field of psychology in that day. As Thomas KUHN has argued (KUHN 1962), it is often the young scholars, the ones who are fortunate enough to arrive on the cusp of change, who benefit most. Since they have not yet been indoctrinated into any particular paradigmatic framework, they are free to adapt to the new and the adventurous. Indeed, I and many of my peers found ourselves unencumbered by behaviorist tenets and free to follow the new winds blowing through the academy. And it was great fun. The revolution was in full swing and it wouldn't be long before we could actually use words like "representation" and "consciousness" and still get our articles published in peer-reviewed journals. A heady time indeed.

The Intellectual Cost of the Revolution

But, as was hinted at earlier, there was a price being paid for all this, and for a long time, no one seemed to notice. Experimental and theoretical psychology, having been largely swept up in the cognitive revolution, took on a particular cast, one that was characterized by: (a) a fascination with the problems of memorial representation while slighting the problem of how knowledge is acquired and, (b) a neglect of the basic DARWINIAN heuristics, the ones that counsel us to question whether proposed models and theories make sense in terms of the adaptationist principles of evolutionary biology.

As these shifts in interest wove their way through our discipline they left their fingerprints behind in the form of some raw but useful numbers that can be used to chart them. The following sections present historical data mapping out the waxing and waning of interest in each of the key topics of *learning, memory,* and *evolution*. They are in the form of

Evolution and Cognition | **105** | 2003, Vol. 9, No. 2

frequency counts that result from a series of sweeps through the large archival base compiled by the American Psychological Association and presented in its PsychINFO files. Details on the manner in which the search was carried out follow.

Search Methods

The frequency counts were obtained by plumbing the depths of the PsychINFO data base using particular key words to define the capture net. The data in Tables 1 and 2 used *learning* for one sweep and *memory* for the other. A similar set of figures for the term *evolution* are presented in Table 3. The full PsychINFO data base was scanned, one decade at a time, with each sweep pegged to a key word. This process yielded the frequencies with which published articles in each decade of the 20th century used these key terms in either their title or abstract.

This technique has some obvious difficulties. I do not believe they compromise the outcomes but they need to be acknowledged. First, there is huge demographic pressure on these frequencies. The sheer number of citations is forced upward as the number of journals published and individuals doing research increased. Hence, the interesting patterns are not the raw figures themselves (which, indeed, almost always show decade by decade increases), but the relative shifts in interest in the target issues.

Second, there are more than a few duplications. The PsychINFO system picks up any article that uses any selected key term. Therefore, any article that used both of the terms *learning* and *memory* in either the title or the abstract would end up being counted in both sweeps. Rather than go through the processes of eliminating duplications (looking at the table will give you can get a feeling for how tough this task would become by the latter part of the century), the decision was made to merely accept these duplications uncorrected. Since the primary concern is with the *shifts* in interest, these overlaps don't seem to present any fatal flaws—in fact, one could argue that they represent true reflections of shared interest.

Third, I wanted these scans to reflect contemporary work as well. Consequently, citations made during the years 2000 to 2003 are included. For consistency, these figures were projected out under the (doubtlessly shaky) assumption that interests in these topics will continue at the same rate for the rest of this decade. The concern is not with the prescient accuracy or lack thereof of this process; the numbers are useful in that they provide a sense of

	Key Word			
Decade	Learning	Memory	L/M Ratio	
1900–1909	41	109	.38	
1910–1919	267	259	1.03	
1920–1929	661	812	.81	
1930–1939	3,087	1,213	2.54	
1940–1949	4,319	1,093	3.95	
1950–1959	7,931	3,365	2.36	
1960–1969	17,920	3,916	4.58	
1970–1979	29,729	11,643	2.55	
1980–1989	35,743	17,376	2.06	
1990–1999	49,072	32,141	1.53	
2000-2009	55,028	38,901	1.41	

Table 1. The number of publications per decade on the key topics of *learning* and *memory* using a full-scan PsychINFO search. The figures for 2000–2009 are projections based on the assumption that the rate of publications for the early years of the current decade will continue throughout it. The L/M ratio is the proportion of articles on *memory* relative to the number published on *learning*.

how popular the topics encompassed by these three key terms are at the present time.

Fourth, as the field of psychology expanded, the horizon of topics included in the data base expanded—and at a stunning rate. Over the last half of the century the number of journals increased dramatically and the range of topics included in the data base expanded in parallel fashion. The raw numbers in Table 1 and the first column of Table 3, which are based on every journal in the capture net (including those in related fields and such semi-scientific publications as *Psychology Today*), show just how stunning this increase was. As a result of this expanded breadth of coverage, the key terms began picking up articles and books that lay outside the primary concerns. For example, *learning* began to capture articles that focused on educational issues, pedagogy, didactic approaches, self-help techniques and similar topics. The term memory began snagging semi-scientific pieces on "how to improve your memory" by any of a number of (highly questionable) techniques and evolution started capturing articles that were more of a political/educational nature than scientific (e.g., debates over the teaching of evolution vs creationism).

Consequently, a second scan was run limiting the search to those journals that had either been around since psychology's earliest days and/or were ones that primarily published formal theoretical articles and research using laboratory-based, empirical methods. In particular, the scan was limited to a select few, highly regarded, intensely peer-reviewed journals. The results from this scan are in Table 2 for

Evolution and Cognition | 106 | 2003, Vol. 9, No. 2

	Key Word			
Decade	Learning	Memory	L/M Ratio	
1900–1909	13	35	.37	
1910–1919	86	83	1.04	
1920–1929	103	117	.88	
1930–1930	768	249	3.08	
1940–1949	888	275	3.23	
1950–1959	1,811	1,154	1.57	
1960–1969	3,400	1,657	2.05	
1970–1979	3,458	1,867	1.85	
1980–1989	2,527	1,852	1.36	
1990–1999	2,913	2,541	1.15	
2000-2009	3,050	3,051	1.00	

Table 2. The number of publications per decade on each of the key topics of *learning* and *memory*. This search was restricted to the historically more "scientific" journals (see text for the full list). The figures for 2000–2009 are projections based on the assumption that the rate of publications for the early years of this decade will continue throughout it. The L/M ratio is the proportion of articles on "memory" relative to the number published on "learning."

learning and *memory* and the right-most column of Table 3 for *evolution*. The journals included in this scan are:

- 1. Journal of Experimental Psychology (including the several spin-off sub-journals established as the need for them grew)
- 2. Quarterly Journal of Experimental Psychology (similarly here)
- 3. Psychological Review
- 4. Psychological Bulletin
- 5. Journal of Comparative and Physiological Psychology, up to the 1980s when it split into:
- 6. Journal of Comparative Psychology and
- 7. Behavioral Neuroscience
- 8. Animal Learning and Behavior (beginning in the 1970s)
- 9. Journal of Human Behavior and Learning (beginning in the 1970s)
- 10. Journal of Verbal Learning and Verbal Behavior, from the 1960s through the 1980s when it became:
- 11. Memory and Language
- 12. Memory and Cognition (beginning in the 1970s)
- 13. Learning and Individual Differences (beginning in the 1980s)
- 14. Learning and Motivation (beginning in the 1970s).

While the issues that are of primary interest here do have considerable overlap, I want to keep them separate for the purposes of exposition. First let's

	Evolution				
Decade	All Journals	"Scientific" Journals			
1900–1909	171	14			
1910–1919	180	21			
1920-1929	908	12			
1930–1939	857	23			
1940–1949	212	6			
1950–1959	271	7			
1960–1969	897	9			
1970–1979	1,466	19			
1980–1989	3,596	44			
1990–1999	7,774	50			
2000-2009	11,009	94			

Table 3.The number of publications per decade in the key area of *evolution*. The column marked "All journals" used a full-scan search (PsychINFO), the other used only those "scientific" journals noted in the text. The figures for 2000–2009 are projections based on the assumption that the rate of publications for the early years of this decade will continue throughout it.

take a look at *learning* and its epistemic partner, *memory*; below we'll examine the patterns of use surrounding the term *evolution*.

Learning and Memory

A look at the early pattern of usage of these terms is instructive. When the empiricist approach to philosophy, led by the loosely confederated group known as the British Empiricists, finally triumphed over Cartesian rationalism with its nativist leanings, a deep obligation was incurred. If you are committed to environmentalism and wish to maintain the *tabula rasa* assumption, even in a diluted form, you have to be able to come up with some explanation of how knowledge gets acquired. Nativism, as has often been noted, is a bit a "copout" in this regard. Nativists needn't worry about how the mind becomes epistemically populated, they simply assume it came fully equipped on arrival.

An empiricist, on the very other hand, needs to entertain some mechanism through which it is possible to understand, at least in principle, how the underlying knowledge base that drives behavior is acquired. Since behaviorists were committed to this brand of empiricism, the topic of learning became central to their enterprise. And, not surprisingly, the parallel topic of memory, which is primarily concerned with the *representation* of knowledge, lost ground. This pattern of shifting fo-

Evolution and Cognition | 107 | 2003, Vol. 9, No. 2

cuses is revealed by the frequency with which these two key terms were used in the decades just before and during the heyday of behaviorism. In fact, there are several intriguing patterns in Tables 1 and 2 that provide a good feel for the waxing and waning of interest in these topics over the full century.

(1) During the first decade *memory* is actually a more popular topic than *learning* although the data base is small. Over the first three decades the number of articles that focused on the two topics is virtually identical, and since these patterns show up in both the broad and the narrow searches they would appear to be reflective of psychology's broad interests during these years.

(2) The first dramatic shift takes place in the 1930s. By this time behaviorism was firmly in place as the dominant approach in psychology. As good descendants of British Empiricism, the practitioners took their obligations on with a vengeance. Throughout the 1930s and 1940s the number of publications that focused on *learning* averaged over three times the number on *memory*. This pattern is seen in both the narrow and the full scans although it is more dramatic in the full scan.

(3) This dominance continued throughout the 1950s, 1960s and, to some extent, into the 1970s. Here the two scans show slightly different patterns. In the unrestricted scan, learning continues to be a more popular topic throughout this period although there is a detectable shift. From the 1940s through the 1960s *learning* is the topic of investigation 3.47 times as frequently as *memory*. In the following decades the ratio undergoes a monotonic decline to its current level of less than 2 to 1. In the restricted scan the abruptness of the shift is a bit more muted. Here, *learning* is cited 2.06 as often as *memory* from the 1940s through to the 1960s but then rapidly loses ground so that the two terms are now essentially on an even footing.

(4) In the restricted scan, the number of articles that focused on *learning* actually drops after the 1970s and still has not completely recovered. Given that the number of journals and researchers increased dramatically during this era, this decline in interest in the topic is diagnostic of a dramatic intellectual shift.

(5) In the last two decades of the century, there is a sense of increased attention paid to *learning*. It is almost certainly the case that this renewal is due to the appearance of connectionism, the first new formal theory of learning to appear in several decades (Rumelhart/McClelland 1986; McClelland/ Rumelhart 1986).

(6) From the 1980s to the current decade, with the cognitivist orientation gradually but effectively taking control, *memory* slowly but inexorably overtakes *learning* as a topic of investigation. These are the decades during which the cognitive revolution matured and became as dominant over the whole of psychology as behaviorism had been during the 1930s and 1940s, if not more so. This modern version of cognitivism embraced and nurtured research on the form and, importantly, the underlying neuroanatomical representation of held knowledge with diminished interest in investigating how it was acquired.

The interpretation of these patterns seems pretty straightforward. The burst of interest in learning in the '30s and '40s was driven by the field's longstanding commitment to Lockean ideals. Memory was important, of course, but it tended to be treated as an adjunct to the more basic topic of *learning*. The late-century shift away from *learning* appears to be a classic case of "throwing out the baby with the bath water." The abandonment of behaviorism led to the systematic neglect of the very topics that had lain close to its conceptual heart and no topic was closer to the core of behaviorism than learning. Admittedly, the failure to continue with the study of learning may have had a lot to do with the manner in which it had been studied in the decades just prior to the emergence of cognitivism. A quick perusal of such volumes as HULL's Essentials of behavior (1951) and A behavior system (1952) or FERSTER and SKIN-NER'S (1957) Schedules of reinforcement will provide a modern psychologist with all he or she would need to understand why the study of learning was largely vacated by cognitivists. Indeed, the lingering concern with learning was driven largely by the influential theory of Robert RESCORLA (1967, 1988; RES-CORLA/WAGNER 1972). RESCORLA's point of view, of course, is distinctly cognitive in nature and allowed those with lingering behaviorist leanings and those who preferred to work with non-human subjects to blend in with the emerging revolution.

Interestingly, the research that began to appear toward the end of behaviorism's dominance explored human functions, particularly those involved in language and thought. Psychologists began dusting off such archaic methods as asking participants what their experiences were, how clear and representative their images were, how they went about solving problems, what words they could recall, how confident they felt in their decisions. At first, these studies were carried out under a behaviorist banner. MILLER's (1951) influential *Lan*-

Evolution and Cognition | 108 | 2003, Vol. 9, No. 2

guage and communication is notable for the line in its preface remarking that "The bias is behavioristic not fanatically behavioristic, but certainly tainted by a preference." One of the first journals to routinely publish studies that followed in this line of research began life in the 1960s as the *Journal of Verbal Learning and Verbal Behavior*. By the 1980s the power of the revolution was felt, *JVLVB* disappeared and was reborn phoenix-like as *Memory and Language*— "learning" was gone, replaced by "memory," "verbal" was gone, replaced by "language." And, "behavior" was merely gone—no replacement needed.

To get a feeling for how this cognitivist Zeitgeist functioned, it is instructive to look in some detail at a topic of investigation that first appeared during this period, implicit learning. Implicit learning is the process whereby knowledge is acquired largely independent of awareness of both the process and the products of learning (REBER 1993). Implicit learning takes place naturally when an individual attends to and works with a complexly structured stimulus domain. It typically results in the induction of a tacit knowledge base that captures many of the structural features of the displays. As a topic of investigation, it is one of the few that emerged early in the cognitive revolution that focused squarely on knowledge acquisition and, because its history overlaps almost perfectly with the shift from behaviorism to cognitivism, it makes for a relevant case study.

A short "aside" on the subject of implicit learning

The first reports appeared toward the end of 1960s (REBER 1967, 1969) just when the move away from behaviorism was going public. On the face of it, this was a topic that should have been embraced by the emerging cognitive revolution. It focused on learning, invited questions about the role of consciousness and had direct implications for developmental psychology, which was undergoing is own process of cognitivization. Moreover, implicit learning maintained links with behaviorism for, by its very nature, the process takes place without the blessings of awareness.

No matter, the topic lay peacefully in the journal pages, virtually unnoticed and uncited for a surprisingly long time. It took nearly three decades on the fringes before it became, quite suddenly, an official "hot topic." Just how "hot" is revealed by the results of the PsychINFO scan in Table 4. The number of publications on the topic jumped from single to triple digits from the 1970s to the 1990s. In addition to this hundred-fold increase in published papers,

Decade	Implicit learning	Implicit memory
1960-1969	2	0
1970–1979	2	0
1980–1989	21	55
1990–1999	244	277
2000-2009	399	914

Table 4. The number of publications per decade on the topics of *implicit learning* and *implicit memory* based on a full-scan search. The figures for 2000–2009 are projections based on the assumption that the rate of publications for the early years of this decade will continue throughout it.

three books were published in rapid succession (Berry/Dienes 1993; CLEEREMANS 1993; REBER 1993), two major journals (*Psychonomic Bulletin & Review* and *Trends in Cognitive Sciences*) featured it as a key topic of research, and in 1997 the "official" imprimatur was granted to the topic with the publication of the *Handbook of implicit learning* (Stadler/Frensch 1997).

This mushrooming of interest, however, was not driven by a sudden concern with mechanisms of acquisition, but by the discovery of the phenomenon of *implicit memory*. As Table 4 also shows, while this flurry of research on *implicit learning* was going on, the topic of *implicit memory* was undergoing a parallel growth. At first blush, this looks like a classic "chicken and egg" problem. Did the sudden interest in *implicit memory* drag along *implicit learning* or did the emerging attention to *implicit learning* pull along memory researchers? For anyone who has worked in this area, the answer is unambiguous: *learning* rode in on the coattails of *memory*.

Researchers who were interested in the unconscious acquisition of knowledge suddenly found their work in the spotlight, not because of any breakthroughs in research or any novel insight. The discovery⁵ of the existence of memorial representations that had causal roles to play on behavior but remained outside the reach of awareness struck a chord among cognitive psychologists. Within a span of ten years, implicit memory went from a nontopic to being one of the most cited key terms in all of psychology. Note the pattern of publications in Table 4. The growth in interest in the role of the unconscious in acquisition only occurred when the topic of memory was recognized as having an implicit component (see SCHACTER 1987). While implicit learning "enjoyed" two decades of existence before *implicit memory* was even recognized as a topic, it was quickly surpassed by it. This asymmetry in interest is even more marked if one looks at the manner in which these two topics are handled in stan-

Evolution and Cognition | 109 | 2003, Vol. 9, No. 2

dard texts in cognitive psychology. Of the baker's dozen modern cognition texts sitting in my book-shelves, every one has a section on *implicit memory* but not one even mentions *implicit learning*.

The Consequences of Favoring Representation over Acquisition

The relative neglect of learning in favor of memory has had two primary consequences. First, it has produced a modern developmental psychology that tends to focus on knowledge representation while neglecting knowledge acquisition. The growth of knowledge and understanding in children has come to be viewed more as a maturational process than one in which environmental circumstances play an important role. Indeed, in much of contemporary developmental discourse the distinction between learning and memory is lost. When it is concluded that some particular concept, word, category or action is part of a child's repertoire it is typically presented as simply that with little or no concern about the underlying mechanisms through which such knowledge was acquired. This movement away from learning was encouraged by the gradual emergence of Piagetian theory. Piaget, while introducing a dialectical framework for understanding how children's representations changed over time (hence his interest in errors that children made), tended to view such changes as the natural progressing of a maturational process.

Second, the neglect of learning was accompanied by an increase in the tacit assumption of genetic deterministic models. In the centuries-long debate between nature and nurture, the cognitive revolution was accompanied by a distinct, inexorable swing of the pendulum back toward nature. There is little doubt that the CHOMSKYAN revolution in linguistics played a significant role here. CHOMSKY famously assumed the existence of a genetically determined "language organ" which held, in some abstract form, the underlying rules of universal grammar and guided the emergence of language in the developing child (CHOMSKY 1966, 1972). CHOMSKY's perspective gave rise to a program of research that culminated in what has become known as *modularity* theory (see FODOR 1983 for the beginnings of this movement and KARMILOFF-SMITH 1992 for a critique of it). Interestingly, while there is no necessary link between nativism and modularity (ELMAN et al. 1997), the movement toward modularity was carried along primarily by those with strong nativist leanings (e.g., PINKER 1994).

This particular trend toward nativism was distinguished from the approach championed by Piaget primarily because, following CHOMSKY, many theorists were assuming a "content" specific endowment rather than a "process" specific one. That is, rather than maintaining that particular processing mechanisms were inborn, the actual epistemic content of mind was assumed to have been genetically programmed in (see REBER 1973 for the nature of this distinction and the role it played in early theory). In PIATELLI-PALMARINI'S (1980) edited volume (which, interestingly, carried the subtitle *The debate between Noam Chomsky and Jean Piaget*), this distinction can be seen in stark relief.

However, this drift toward innateness was a most unusual one for it took place virtually completely without the blessings of evolutionary biologists. In fact, the movement has an oddly paradoxical element in that while its proponents made bold assumptions about genetically encoded mechanisms they were, throughout their decades of maximal influence, singularly uninterested in even discussing DARWINIAN adaptationist issues. And, in this they mirror the second of the major consequences of the cognitive revolution, the diminution of the role of adaptationist thinking.

Darwinian adaptationism

As with the shifting balance between *learning* and *memory*, the best way to get a sense of how the interest in Darwinism changed over the last century is to look at the citation data. Table 3 shows the output of the same broad and narrow scans of the PsychINFO data base, this time using the key word evolution. The patterns here share some features with the earlier scans but also have some unique elements. First, the topic of evolution was quite popular early on. In fact, it was as frequently the focus of scientific work as either *learning* or *memory* up through the end of the 1920s. Interestingly, of the three topics it was considerably more frequently cited in the larger data base than in the narrower, as can be seen by comparing the raw numbers for all three terms during this period. Whereas *learning* and *memory* were the topics of publications roughly three times as often in the full scan than in the narrow, with evolu*tion* this ratio is closer to 10 to 1.

The burst in interest that appears in the 1920s is at least partly due to the emerging grand synthesis in evolutionary biology that developed when Mendelian principles were blended with DARWINIAN theory. The growth in the 1930s reflects this continu-

Evolution and Cognition | 110 | 2003, Vol. 9, No. 2

ing interest along with encouragement from behaviorists who treated evolution as legitimizing their dependence on cross-species generalizations and were deeply sympathetic with the ethology movement that was emerging among European zoologists (e.g. LORENZ 1937).

The next two decades, however, show a precipitous decline in interest. In fact, evolution as a key term almost disappears from the scientific forum. In the 30 years from 1940 until 1969, the narrow scan yields a grand total of only 22 publications using evolution as a key term. Even in the broad scan, the number of publications drops by an astonishing 75% from the decade of the 1930s to the 1940s. The gradual demise of behaviorism may have played a role here, but it is unlikely to have been the critical element. And, while the emergence of cognitivism certainly worked to delay the resurgence in adaptationist thought until the 1990s, it's difficult to see how it could have had much impact in the post war era. Rather, the virtual abandonment of Darwinism during this period was primarily a reaction to an ever-deepening suspicion of adaptationist thinking—and for several good reasons.

First, psychology was still struggling with its long and occasionally ugly history of exploiting adaptationism to draw invidious distinctions among peoples of differing ethnic and racial ancestries. The excesses of the eugenics movement and a recognition of the impact it had on immigration, education, health care, intelligence testing, and social policy (see HOTHERSALL 1995) was a compelling intellectual brake—especially when all of it was brought into horrific clarity by Nazism. And just when we began to hope we were over such embarrassments as GAL-TON'S (1874, 1880) and GODDARD'S (1912, 1917) arguments, works such as those authored by RUSHTON (1988) and HERRNSTEIN and MURRAY (1994) surfaced to show us again how treacherous this approach can be. The angry and combative reception that WILson's book Sociobiology (WILSON 1975) received is indicative of the strong feelings held by many (see, e.g., LEWONTIN/ROSE/KAMIN 1984).

Second, evolutionary biologists such GOULD and LEWONTIN pointed out that the line between adaptation and exaptation was often a fine one. In their seminal "spandrels" paper (Gould/LEWONTIN 1979), they cautioned against drifting into a pan-adaptationism where all surviving organic features and behavioral propensities were assumed to have evolved because of particular adaptive roles in the species' past. Many adaptive forms and functions, they cautioned, exist today, not because they necessarily possessed singular utility, but because they "piggybacked" on other related forms and functions or were *spandrels*—that is, necessary consequence of the emergence of related structures.

The result of these several forces was that Darwinism, so long the intellectual umbrella under which psychology flourished, was shunted aside like an embarrassing cousin. As the years went by it became increasingly clear that this was no small debt that was accumulating. Since its emergence as a distinct discipline, psychology had been codified, structured, and embodied within a functionalist's frame of reference. DARWINISM had been the initial binding force that brought together the diverse elements of our science. Without DARWIN there would have been no basis for cross-species studies, there would have been no ethology, no physiological psychology, no comparative investigations. Individual differences would have been viewed as they were in HELMHOLTZ's day, as annoying sources of error variance. Fields such testing and evaluation, assessment, diagnostics, indeed most of modern psychometrics, would have been reduced to little more than desperate attempts to declutter our data base. The mind itself would have been viewed as Titchener had naively counseled, an object whose structure was to be unpacked independent of any considerations of its functions. Emotions would have been seen as a compendium of modes of expression rather than as exquisitely evolved adaptations to the vagaries of an ever-changing environmentand no one would ever have wondered whether or why a nonhuman primate could behave deceptively (SEYFARTH/CHENEY 2002).

The cognitive revolution, for all it did to free psychology from the shackles of behaviorism, actually worked against the emergence of adaptationist thinking. Indeed, it assisted in trussing up the field in a very different conceptual straightjacket. Enamored of information processing models, of computer simulations, of those clever theories in which so many boxes danced across the pages of journals and books, we found ourselves forgetting to ask the most basic questions: Do these models of mind make any DAR-WINIAN sense? Could something like a language acquisition device, fully equipped with the contents of Universal Grammar have evolved within the time frame set out for it? Do people really shunt information from one kind of memory store to another? Did tightly encapsulated modules have any legitimate ontological status-did they have the requisite adaptive characteristics? Could the mind/brain really have boxes labeled "input," "sensory buffer," "push-

Evolution and Cognition | 111 | 2003, Vol. 9, No. 2

down stack?" Did feature-detecting demons shout to each other about aspects of a stimulus display? Did sensory systems really pick up n + 1 stimulus elements in each conceptual glance?

Of course, operating within the confines of conceptual models is a time-honored technique; we are used to theories that say, in essence, "the mind operates just like..." But the early decades of the cognitive revolution were marked by the development of models of mind that began to fudge the simulation– emulation boundary and sought to capture the data base without ever invoking the classic DARWINIAN heuristic: "Could systems that functioned like these models implied possibly have evolved and could they have done so in the time frame specified?" Viewed from an adaptationist's perspective it became clear that, in Gerald EDELMAN's famous remark, much of this early work "wasn't even wrong."

Intriguingly, as noted above, this era was marked by an odd and almost paradoxical sequence of developments. On one hand, many of the new and influential cognitive scientists followed the CHOM-SKYAN move and adopted a strongly nativist perspective. Nativism is, virtually by definition, bound up with genetics and genetics is, equally compellingly, linked to principles of evolutionary biology. Yet the new nativists were singularly unwilling to take into consideration DARWINIAN, adaptationist principles. With rare exception the question of how, why, and under what circumstances these assumed innate mechanisms evolved was never even asked, let alone debated.

Clearly, the situation is changing, if slowly. The citation frequencies for the 1990 and those projected through the next decade suggest that Darwinism has most certainly re-emerged. For the most part this resurgence is due to the growth in the discipline of evolutionary psychology. Basic processes like learning and conditioning are now understood as being modulated by mechanisms with adaptationist elements. Social processes are now appreciated as affected by species-specific predispositions and such basic operations as in-group cooperation, mate selection, the emergence of consciousness, the development of aesthetic

judgment, altruism, group selection pressures, and why patterns of prejudice and distrust tend to erupt under particular conditions of environmental stress are now being viewed within DARWIN-IAN contexts.

Sometimes the game is easy to play and non-controversial. The gradual downward shift in the position of the larynx was shown to have adaptationist properties since it led to the creation of the supralaryngeal vocal tract allowing humans the freedom to produce the rich phonology of human speech and, of course, to choke on their food (LIEBERMAN 1984). At other times things have gotten a bit sticky, as when various scenarios were sketched showing how and why various forms of social interaction might have given rise to a genetically encoded tendency for altruism (RUSHTON 1989), what kinds of biological determinants could lead to suicide (DECATAN-ZARO 1980), where the genetic roots of empathy can be found (PRESTON/DE WAAL 2002) or, perhaps the most difficult problem of all, what kinds of primordial scenes might have encouraged the development of a phenomenologically poignant, self-referencing consciousness (FLANAGAN 1991, Chapter 8; REBER 1997).

But, no matter. All sciences face problems when doing "normal science." The adaptationist's method, prone though it is to the "just-so story" fallacy, and open as it is to abuse by the self-serving, is the generally accepted way of doing evolutionary psychology. Indeed, as de Waal proclaimed in a recent paper, "evolutionary approaches are on the rise ... and have the potential to bring an all-encompassing conceptual framework to the study of human behavior" (de Waal 2002, p187).

Independent of the accuracy of de Waal's prescience, the citation data certainly suggest that there is a gradual redressing of the neglect of the issues of acquisition and adaptationism. Closing these intellectual lacunae will certainly enhance the emergence of a mature psychology. The hope is that future work will take place in the context of a science that routinely invokes basic heuristics when entertaining models of mind, theories of behavior, or metaphors for the human brain. Specifically,

(1) When regular and consistent patterns of behavior are observed, toy with models of acquisition before assuming a nativist ontology. Learning almost certainly plays a larger role in human func-

Author's address

Arthur S. Reber, Department of Psychology, Brooklyn College of CUNY, Brooklyn, New York, 11210, USA. Email: areber@brooklyn.cuny.edu. tioning than is currently believed by most psychologists.

(2) Look for the adaptive functions of hypothesized mechanisms and question whether they appear to be intrinsically linked with particular structures and forms or

Evolution and Cognition | **112** | 2003, Vol. 9, No. 2

whether they might be exaptively associated with parallel structures and forms.

(3) Consider whether and how a system with these properties could have evolved within the time frame indicated.

(4) Look for the entailments of hypothesized mechanisms and try to determine if their presumed structures do violence to what is already known or suspected about human behavior and the underlying neural systems that mediate them.

(5) Question whether the specific entailments of such a mechanism make sense in light of the accepted models of evolutionary biology.

If we hold dear to these heuristic devices we should be fine in the coming decades.

Author Note

Many of the issues discussed in this paper were first presented in the Keynote Address at the inauguration of the New England Institute for Cognitive Science and Evolutionary Psychology, November 2, 2001, Portland, Maine. Special thanks to Laraine McDoNOUGH for her helpful comments and gentle criticisms and to Roberta MATTHEWS for the time to complete this project.

The paper was prepared while the author was supported by Grant #0113025 from the NSF and by a partial fellowship from the Wolfe Institute for the Humanities at Brooklyn College.

Notes

- 1 BAARS's book was an "early" history in which he searched for the roots of the intellectual revolution as it was taking place. His vision of the evolving paradigm was expressed by the areas that he focused on and, of course, by those that he neglected. The two slighted orientations that were to become important were the neurocognitive sciences and the delicate interlacing of cognitive-style approaches with social psychology. Neither absence is terribly surprising. Social psychology had always been much more cognitive than other areas so the shift there was not as dramatic and the neurocognitive aspects were awaiting the development of sophisticated scanning and recording techniques that opened up areas of research that had only been imagined in decades past.
- **2** The case of forensic psychology was complicated by the fall from grace of its most ardent promoter, Hugo MÜN-STERBERG. MÜNSTERBERG, a German immigrant and supporter of conciliation and peace with Germany, argued strongly and publicly for his point of view during the years leading up to the first World War. Sternly criticized for this stance by the media, many politicians and even his colleagues at Harvard, his research program focusing on psychology and the law fell into a vague form of benign neglect that, in all likelihood, would have been its fate

even without the behaviorist intellectual hegemony that was to follow.

- **3** CHOMSKY also championed an updated version of Cartesian nativism in his hypothesizing of an innately given, content-specific, universal grammar as the biological core of all human languages. Paradoxically, CHOMSKY, despite putting such theoretical weight on genetic and biological factors, has, until very recently (HAUSER/CHOMSKY/FITCH 2002), been reluctant to even discuss issues of evolutionary biology. I'll have more to say on this discoordination later. It is of considerable importance in understanding the nature of the emergence of cognitivism.
- **4** Many are unaware that one of psychology's Nobel Prize winners, Herbert SIMON, was originally trained in political science. When he went to the then Carnegie Institute of Technology it was as an instructor in political science and among the first courses he offered were ones in constitutional law. Only later did his interests take him into economics, computer sciences, and eventually psychology.
- **5** Or, more accurately, the rediscovery. Such prominent neurologists as KORSAKOFF (1889), CLAPARÈDE (1951), BREUER, JANET, FREUD and PRINCE (see PERRY/LAURENCE 1984) had all reported patients with various forms of amnesia who showed behavioral evidence of memories of events in the absence of awareness. See SCHACTER (1987) for a overview of this earlier work.

References

- Arnold, M. B. (1945) Physiological differentiation of emotional states. Psychological Review 52: 35–48.
- Arnold, M. B. (1970) Feelings and Emotions. Academic Press: New York.
- Baars, B. J. (1986) The cognitive revolution in psychology. Guilford Press: New York.
- Bartlett, F. C. (1932) Remembering. Cambridge University Press: Cambridge.
- **Beach, F. A. (1945)** Current concepts of play in animals. American Naturalist 79: 523–541.

Bloomfield, L. (1933) Language. Holt: New York.

- **Broadbent, D. E. (1958)** Perception and Communication. Pergamon Press: New York.
- Bruner, J. S./Goodnow, J. J./Austin, G. A. (1956) A Study of Thinking. Wiley: New York.
- Bush, R. R./Mosteller, F. (1951) A mathematical model for simple learning. Psychological Review 58: 313–323.
- Bush, R. R./Mosteller, F. (1955) Stochastic Models for Learning. Wiley: New York.
- **Cherry, E. C. (1953)** Some experiments on the recognition of speech, with one and with two ears. Journal of the American Acoustical Society 25: 975–979.

Evolution and Cognition | 113 | 2003, Vol. 9, No. 2

Cherry, E. C. (1957) On Human Communication. Wiley: New York.

- Chomsky, N. (1959) Review of Skinner's Verbal Behavior. Language 35: 26–58.
- Chomsky, N. (1966) Cartesian Linguistics. Harper & Row: New York.
- Chomsky, N. (1972) Language and Mind. Harcourt and Brace: New York.
- Claparède, E. (1951) Recognition and "me-ness." In: Rapaport, D. (ed) Organization and Pathology of Thought. Columbia University Press: New York, pp. 58–75. Reprinted from Claparède, E. (1911) Reconnaissance et moiité. Archives de Psychologie 11: 79–90.
- **Cleeremans, A. (1993)** Mechanisms of Implicit Learning. MIT Press: Cambridge MA.
- **DeCatanzaro**, **D.** (1980) Human suicide: A biological perspective. Behavioral and Brain Sciences 3: 265–290.
- Elman, J. L./Bates, E. A./Johnson, M. H./Karmiloff-Smith, A./ Parisi, D./Plunkett, K. (1997) Rethinking Innateness: A Connectionist Perspective on Development. MIT Press: Cambridge MA.
- Estes, W. K. (1959) The statistical approach to learning theory. In: Koch, S. (ed) Psychology: The Study of a Science, Volume 2. McGraw-Hill: New York, pp. 380–491
- Fagan, R. (1981) Animal Play Bhavior. Oxford University Press: New York.
- Ferster, C. B./Skinner, B. F. (1957) Schedules of Reinforcement. Appleton-Century Crofts: New York.
- Flanagan, O. (1991) The Science of Mind (2nd edition). MIT Press: Cambridge MA.
- Fodor, J. A. (1983) The Modularity of Mind. MIT Press: Cambridge MA.
- Galton, F. (1874) English Men of Science: Their Nature and Nurture. Macmillan: London.
- Galton, F. (1880) Hereditary Genius (2nd American edition). New York: D. Appleton. Originally published in 1869.
- Goddard, H. H. (1912) The Kallikak Family: A study in the Heredity of Feeble-Mindedness. Macmillan: New York.
- Goddard, H. H. (1917) Mental tests and the immigrant. Journal of Delinquency 2: 243–277.
- Gould, S. J./Lewontin, R. C. (1979) The Spandrels of San Marco and the Panglossian paradigm. Royal Society of London: Proceedings B, 205: 581–598.
- Groos, K. (1898) The Play of Animals. Translated by E. L. Baldwin. Appleton & Co: New York.
- Groos, K. (1899) Die Spiele der Menschen [The Play of Men]. Fischer Verlag: Jena.
- Hauser, M. D./Chomsky, N./Fitch, W. T. (2002) The faculty of language: What is it, who has it, and how did it evolve? Science 298: 1565–1579.
- Herrnstein, R. J./Murray, C. (1994) The Bell Curve: Intelligence and Class Structure in American Life. The Free Press: New York.
- Hull, C. L. (1943) Principles of Behavior. Appleton-Century-Crofts: New York.
- Hull, C. L. (1951) Essentials of Behavior. Yale University Press: New Haven CT.
- Hull, C. L. (1952) A Behavior System. Yale University Press: New Haven CT.
- Karmiloff-Smith, A. (1992) Beyond Modularity: A developmental perspective on cognitive science. MIT Press: Cambridge MA.
- Kassin, S. (2003) Psychology (4th edition). Prentice Hall: Upper Saddle River NJ.
- **Koffka, K. (1935)** Principles of Gestalt Psychology. Harcourt Brace: New York.

- Köhler, W. (1925) The mentality of apes. Translated by E. Winter. Harcourt Brace: New York. German original published in 1917.
- Korsakoff, S. S. (1889) Etude médico-psychologique sur une forme des maladies de la mémoire [Medical–Psychological Study of a Form of Memory Disorder]. Revue Philosophique 28: 501–530.
- Kuhn, T. S. (1962) The Structure of Scientific Revolutions. University of Chicago Press: Chicago.
- Lashley, K. S. (1951) The problem of serial order in behavior. In: Jeffress, L. S. (ed) Cerebral Mechanisms in Behavior. Wiley: New York.
- Lewontin, R. C./Rose, S./Kamin, L. J. (1984) Not in Our Genes: Biology, Ideology and Human Nature. Random House: New York.
- Lieberman, P. (1984) The Biology and Evolution of Language. Harvard University Press: Cambridge MA.
- Lorenz, K. (1937) Imprinting. The Auk 54: 245-273.
- Luce, R. D. (1959) Individual Choice Behavior. Wiley: New York.
- McDougall, W. (1932) The Energies of Men: A Study of the Fundamentals of Dynamic Psychology. Methuen and Co: London.
- Miller, G. A. (1951) Language and Communication. New York: McGraw-Hill.
- Miller, G. A. (1956) Human memory and the storage of information. IRE Transactions on Information Theory IT-2(3): 128–137.
- Miller, G. A. (2003) The cognitive revolution: A historical perspective. Trends in Cognitive Sciences 7: 141–144.
- Miller, G. A./Galanter, E./Pribram, K. H. (1960) Plans and the Structure of Behavior. Holt, Rhinehart & Winston: New York.
- Morgan, C. L. (1894) An Introduction to Comparative psychology. Walter Scott Ltd: London.
- Neisser, U. (1967) Cognitive Psychology. Appleton: New York.
- **Osgood, C. E./Sebeok, T. A. (1965)** Psycholinguistics. Indiana University Press: Bloomington IN. Originally published in 1954.
- Perry, C./Laurence, J. R. (1984) Mental processing outside of awareness: The contributions of Freud and Janet. In: Bowers, K. S./Meichenbaum, D. (eds) The Unconscious Reconsidered. Wiley: New York, pp. 9–48.
- **Piatelli-Palmarini, M. (ed) (1980)** Language and Learning: The Debate between Noam Chomsky and Jean Piaget. Harvard University Press: Cambridge MA.
- Pinker, S. (1994) The Language Instinct: How the Mind Creates Language. William Morrow: New York.
- Preston, S. D./van de Waal, F. B. M. (2002) Empathy: Its ultimate and proximate bases. Behavioral and Brain Sciences 25: 1–72.
- **Reber, A. S. (1967)** Implicit learning of artificial grammars. Journal of Verbal Learning and Verbal Behavior 6: 855– 863.
- **Reber. A. S. (1969)** Transfer of syntactic structure in synthetic languages. Journal of Experimental Psychology 81:115–119.
- **Reber, A. S. (1973)** On psycho-linguistic paradigms. Journal of Psycholinguistic Research 2:289–319.
- **Reber, A. S. (1993)** Implicit Learning and Tacit Knowledge: An Essay on the Cognitive Unconscious. Oxford University Press: New York.
- Reber, A. S. (1997) Caterpillars and consciousness. Philosophical Psychology 10: 437–450.

Rescorla, R. A. (1967) Pavlovian conditioning and its proper

Evolution and Cognition | 114 | 2003, Vol. 9, No. 2

control procedures. Psychological Review 74: 71-80.

- **Rescorla, R. A. (1988)** Pavlovian conditioning: It's not what you think it is. American Psychologist 43: 151–160.
- Rescorla, R. A./Wagner, A. R. (1972) A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In: Black, A. H./Prokasy, W. F. (eds) Classical Conditioning II: Current Research and Theory. Appleton-Century-Crofts: New York, pp. 64–99.
- **Romanes, G. J. (1965)** On comparative psychology. In: Herrnstein, R. J./Boring, E. G. (eds) A Source Book in the History of Psychology. Harvard University Press: Cambridge MA. Originally published in 1882.
- Rushton, J. P. (1988) Race differences in behaviour: A review and evolutionary analysis. Personality and Individual Differences 9: 1009–1024.
- **Rushton, J. P. (1989)** Genetic similarity, human altruism, and group selection. Behavioral and Brain Sciences 12: 503–559.
- Schacter, D. L. (1987) Implicit memory: History and current status. Journal of Experimental Psychology: Learning, Memory and Cognition 13: 501–518.
- Schlosberg, H. (1947) The concept of play. Psychological Review 54: 229–231.
- Seyfarth, R. M./Cheney, D. L. (2002) The structure of social knowledge in monkeys. In: Bekoff, M./Allen, C./

Burghardt, G. M. (eds) The Cognitive Animal: Empirical and Theoretical Perspectives on Animal Cognition. MIT Press: Cambridge MA, pp. 379–384.

- Shannon, C. E. (1948) A mathematical theory of communication. Bell Systems Technical Journal 27: 623–656.
- Shannon, C. E. (1951) Prediction and entropy of printed English. Bell Systems Technical Journal 30: 50–64.
- Shannon, C. E./Weaver, W. (1951) The Mathematical Theory of Communication. University of Illinois Press: Urbana IL.
- Skinner, B. F. (1957) Verbal Behavior. Appleton-Century-Crofts: New York.
- Swets, J. A./Tanner, W. P./Birdsall, T. G. (1961) Decision processes in perception. Psychological Review 68: 301–340.
- Tolman, E. C. (1926) A behaviorist theory of ideas. Psychological Review 29: 352–369.
- Tolman, E. C. (1932) Purposive Behavior in Animals and Men. Century: New York.
- Tolman, E. C. (1948) Cognitive maps in rats and men. Psychological Review 55: 189–208.
- Torgerson, W. S. (1958) Theory and Methods of Scaling. Wiley: New York.
- Watson, J. B./McDougall, W. (1928) The Battle of Behaviorism. Kegan Paul: London.
- Wilson, E. O. (1975) Sociobiology. Harvard University Press: Cambridge MA.

The Effect of Fertility Risk on Relationship Scrutiny

or most people, hav**h** ing a satisfactory romantic relationship determines to a great extent their level of satisfaction with life and their subjective well-being (CAMP-**BELL/SEDIKIDES/BOSSON** 1994; **RUSSELL/WELLS** 1994). A number of studies have shown that problems in marriages and marriage-like relationships may have severe consequences on one's physical and/or mental health (MAYNE et al. 1997; WICKRAMA/ LORENZ/CONGER/ELDER 1997). When compared

to persons with marital difficulties, those in happy marriages were found to have a healthier immune system (KIECOLT-GLASER/FISHER/OGROCKI 1987). In addition, children who grow up with a satisfactory and stable parental relationship appear to go on to encounter fewer difficulties in their later development (FELDMAN/FISHER/SEITEL 1997).

Considering the relevance of close relationships for psychological and physical wellbeing, individuals would be expected to engage, under some circumstances, in relationship evaluation. Under such circumstances, individuals would attempt to break down their notion of a close relationship into its constituent features, examine the prevalence (i.e., presence or absence) of those features in their own relationship, and weigh the degree to which the prevalence of those features contributes to their dissatisfaction or satisfaction with the relationship.

HASSEBRAUCK and SEDIKIDES (under review) label this thorough and systematic processing of relationship-relevant information *relationship scrutiny*. In a

Abstract

This study examined the influence of women's menstrual cycles on romantic relationship scrutinizing, i.e., the degree of systematic processing of relationship-relevant information. In making a global evaluation of relationship satisfaction, 87 women were asked to evaluate their relationship according to 64 specific features of the prototype of relationship quality (HASSEBRAUCK/FEHR 2002). Normally ovulating women (not taking a contraceptive pill) processed information concerning their relationship more systematically during the period of highest fertility. Results were discussed within an evolutionary framework.

Key words

Menstrual cycle, information processing, fertility, relationships.

series of experimental and correlational studies, the authors analyzed the conditions which may have an impact on scrutinizing. Among other findings, they found for example that persons who were dissatisfied with their relationship processed relationship-relevant information more systematically than satisfied persons did. In addition, participants who were more committed scrutinized their relationships more.

A large number of theoretical and empirical

studies show that pair bonding has different consequences for men and for women (see TRIVERS 1972; BUSS 1994; ELLIS 1992). The human species invests a great deal, in terms of biology, time and energy, in their offspring; although in general, women invest more of the above-mentioned investment criteria (TRIVERS 1972). Wherever high costs are involved, the risk of poor investment or deceit is also high. In order to minimize this risk, sex-specific mate selection strategies have developed in the history of human evolution, so-called psychological programs (ALLGEIER/WIEDERMAN 1994; BUSS 1994).

Since theoretically speaking it is women who risk conception with every sexual encounter—unlike men, who may be free of further involvement after a sexual act—they need to be more selective than men in the choice of their partner. Women should keep an open eye for signs that denote the partner's dependability, determination, and his potential for possible long-term investment in a relationship (BUSS/SCHMITT 1993; HATFIELD/SPRECHER 1995;

Evolution and Cognition | 116 | 2003, Vol. 9, No. 2

SYMONS 1979; TRIVERS 1972). Numerous studies have in fact shown that women in general not only tend to be more selective than men (BOTWIN/BUSS/SHACK-ELFORD 1997), but also tend to focus on different characteristics than men when selecting their partner (BUSS 1989). Women also seem to be more realistic with regard to their relationships. SPRECHER/ METTS (1989) for example, found that, contrary to common beliefs, men scored higher on a romantic belief scale than women. Previously, HOBART (1958) and KEPHART (1967) reported that men were more likely to have beliefs such as love will conquer all, and that one should marry for love.

The fact that women are more realistic and are more fine-tuned to the ups and downs in their relationship, is also supported by a number of studies demonstrating that women are "better barometers" of their relationship. In a longitudinal study by KUR-DEK (1993), relationship stability was better predicted with the use of women's rather than men's data. RUVOLO/VEROFF (1997) further add that the women's perceived real-ideal discrepancies within a relationship correlate with marital well-being a year later. In accordance with the results reported above, no significant correlations were obtained when computations were carried out for men. KURDEK (2002), however, did not find that information obtained from wives was sufficient to predict relationship outcomes. This may be due to the fact that this study attempted to predict marital stability. Stability is by definition a variable which is influenced by both partners' decisions. If a person is trying to maintain his or her relationship but the other partner has the desire to end it, the result has to be accepted. Relationship satisfaction, however, is generally more strongly related to wives' evaluations of relationship features than men's.

In the evolutionary past of women, the risk of poor investment was especially high during their fertile phases. Choosing a mate with poor qualities, e.g., a man who was not willing to invest in offspring, would be especially harmful if the woman became pregnant. There is now considerable empirical evidence that women's menstrual cycle effects sexual behavior as well as cognitive processes (see GANGESTAD/COUSINS 2002, for a review). Normally ovulating women (not taking the pill or other hormone-based contraceptives) report an increased sexual desire during their fertile phases (REGAN 1996). They engage in more extra-pair mating, and have more sexual fantasies (BAKER/BELLIS 1995, GANGESTAD/THORNHILL/GARVER 2002). There is also preliminary evidence that women become more discriminating when evaluating potential sex partners during their fertile periods (cf. GANGESTAD/ COUSINS 2002, p171). THORNHILL/GANGESTAD (1999) have conducted an impressive study, in which they demonstrated that women's olfactory preference for the scent of symmetrical men varies across their menstrual cycle. Women's olfactory preference for symmetrical men was greatest during the period of highest fertility across their cycle, provided of course that they were not taking the pill or any other hormone-based contraceptive. PENTON-VOAK/ PERRETT (2000) found that women's preferences for male faces changed during the menstrual cycle. Women in the high fertile phase of their cycle significantly more often chose masculine faces than women in the low fertile phase.

In the face of this literature, I have examined possible consequences of the menstrual cycle on women's scrutinizing. I expect to find that women are more careful during high-fertility phases and inspect their partner and relationship more thoroughly and systematically than women in phases of low fertility. Thus the central hypothesis of the present study is:

Women scrutinize their relationship more during their period of high fertility than women in low fertility phases of the cycle.

Among other reasons, it is therefore important and relevant to investigate this hypothesis because, to my knowledge, no empirical data exists to date to suggest that, dependent on their menstruation cycle, women are varyingly discriminative in regard to their individual relationship and partner. Confirmation of this hypothesis would support the assumption that scrutinizing has indeed developed out of a specific adaptation mechanism that leads to thorough relationship assessment, particularly whenever the relationship itself is in danger, or whenever there is an increased risk of poor investment, such as, for example, during the period of highest fertility.

Method

Participants were 87 women between 18 and 43 years (M = 26.5 years, SD = 6.56) who were involved in a heterosexual relationship at the time of the study. Twenty-two of them were married, 44 were cohabitating, 21 were not living together. Mean relationship duration was 5.3 years (SD = 5.52). They were given 5 Euros for their participation (approximately 5 US\$). A questionnaire to be completed at home was distributed to women who were inter-

Evolution and Cognition | **117** | 2003, Vol. 9, No. 2

ested in participating in the study. Participants were instructed not to discuss any aspects of the questionnaire with their partner prior to completing it and to put it in an enclosed envelope and send it back after completion.

The questionnaire included items assessing perceived relationship satisfaction, which was measured using a German version (HASSEBRAUCK 1991) of HENDRICK' (1988) seven-item Relationship Assessment Scale (RAS). This scale is a one-factorial reliable instrument in measuring global relationship satisfaction (alpha = .92).

Following this, 64 features of the prototype of relationship quality (HASSEBRAUCK/FEHR 2002) were included. Participants were asked to use a 7-item rating scale to determine the extent to which these features were present in their relationship, ranging from 1 "not at all present in my relationship" to 7 "very much present in my relationship". Trust, displaying emotions, sexual satisfaction, having and allowing for freedom, etc., are typical features of the prototype of relationship quality. (For a detailed description of these features refer to Table 1.) Participants use these features when they are evaluating how satisfied they are with their relationship (HAS-SEBRAUCK/ARON 2001). The 64 features of relationship quality demonstrate high internal consistency (*alpha* = .97). Finally, the participants were asked (1) whether they were taking the pill or any other hormone-based contraceptives, (2) to state the number of days since the start of their last menstrual cycle (they were provided with a calendar to assist with this task), and (3) to provide details of the typical length in days of their menstrual cycle, provided that they have a regular cycle.

Operationalization of Scrutinizing: Relationship scrutinizing is a matter of information processing. Extensive literature in the field of social cognition indicates that a number of factors, such as mood, personal relevance, and cognitive load, just to mention a few, seem to induce either a systematic or a heuristic mode of information processing (CLORE/ SCHWARZ/CONWAY 1994). Scrutinizing refers to the systematic processing of relationship relevant information. When individuals are asked to rate how satisfied they are with their relationship, they may base their global evaluative judgement on a thorough and careful analysis of all aspects that, in their eyes, are relevant for a good relationship. In other words, they may compare their momentary relationship with the prototype of a good relationship, which is their mental point of reference of relationship quality and thus, their standard of assessment (HASSEBRAUCK/ARON 2001). These individuals are assumed to scrutinize their relationship. Individuals may just as well avoid thoroughly processing information and rely instead on heuristics and peripheral cues as a basis of their relationship assessment, e.g., their present mood (SCHWARZ/BLESS 1991) or the fact that the marital partner failed to tidy the breakfast table in the morning.

In other words, the *correlation* between the *global* evaluation of relationship satisfaction and the intensity of specific features of the prototype of relationship quality is an indicator of the degree of scrutinizing engaged in. Individuals who tend to engage in more scrutinizing than others should have higher correlations between the features of relationship quality on the one hand, and relationship satisfaction on the other hand, than those who engage in less scrutinizing. Scrutinizing is, as such, the match between one's global relationship judgment and the perceived features of the prototype of relationship quality. In other words, the statistical relation (i.e., the correlation) between overall relationship satisfaction and judgments of specific relationship quality features is an indication of the degree of systematic information processing. The stronger the correlation, the higher the level of scrutiny.

Definition of high fertility and low fertility subgroups: Forty-eight, of the 87 participants reported taking the pill. The remaining 39, in line with THORNHILL and GANGESTAD (1999), were rated as low or high in terms of fertility risk. I determined the day of the menstrual cycle the women were at (forward method) and, as suggested by THORNHILL and GANG-ESTAD, the women's reported cycle length (M = 28.0days, SD = 1.40, range = 24–30). Women who have longer cycles ovulate on average later in their cycle than women who have shorter cycles. In the same manner as Thornhill and Gangestad, I assumed that the typical day of ovulation is 15 days prior to the end of their typical cycle (for example, on day 15 in a 30day cycle). Only during days 6 to 14 of the menstrual cycle is fertility risk at least .15 (THORNHILL/ GANGESTAD 1999, p182). In the remaining context, I will be referring to this group as the high fertility risk group (HFR, N = 19). When considering women's reported cycle length (backward method), participants could be classified into different groups, as when using the forward method which assumes an average 28 day cycle. In such cases, I decided to group participants according to their reported cycle length. Although women taking hormone-based contraceptives face a low risk of

Evolution and Cognition | 118 | 2003, Vol. 9, No. 2

conception, these women may differ from women not taking the pill in important ways (e.g., their sexual attitudes, their commitment to the relationship). Thus, I decided to exclude these women from the low fertility risk group (LFR, N = 20).

Results

To begin with, I calculated the statistical association between the relationship satisfaction composite and a mean composite of the 64 relationship quality features for each group. For the high fertility risk group, reliance on specific relationship quality features accounted for approximately 88% of the variance (r = .94) in relationship satisfaction. For the low fertility risk group, however, reliance on relationship quality features accounted for only 61% of that variance (r = .78). This difference is significant, z = 2.06, p < .05, and cannot be traced back to disparities in relationship satisfaction between the two groups, t(37) = 1.14, p > .26, or to discrepancies in variances F(1,37) = 2.93, p > .095.

I then proceeded to examine single-order correlations between the 64 features of the prototype of relationship quality and relationship satisfaction for the HFR group and the LFR group independently. Fifty-two of the 64 correlations were higher in the HFR than in the LFR group. These results are all the more impressive given that participants, as a whole, were fairly satisfied with their relationships (M = 5.66, SD = 1.09), thereby attaining the maximum possible correlation. (Since these features are highly correlated, it is not appropriate to use the sign-test to check whether this distribution differs significantly from chance.)

Taking a closer look at each individual correlation, I found that women in the HFR group, unlike the LFR group, base their relationship satisfaction more on features such as "being there for each other", "knowing partner", "security", "sexuality", and "safety", for example, to mention a few (see Table 1, for details). This pattern provides support to the assumption that highly fertile women are especially responsive to those qualities of their partners and relationships which indicate stable and secure pair bonding. Women in the LFR group, however, base their relationship evaluation more on features dealing with similarities: "mutual goals", "similar interests", "similar beliefs", features which are considered peripheral for the prototype of relationship quality (HASSEBRAUCK 1997).

In order to analyze these differences in content even further, I calculated 4 sub-scales of the 64 features of the prototype of relationship quality, which correspond to the four factors-intimacy (alpha = .87), agreement (alpha = .82), independence (alpha = .82), and sexuality (alpha = .84), reported by HASSEBRAUCK and FEHR (2002). HASSEBRAUCK/FEHR (2002, Study 4) have already determined that the correlation between intimacy and relationship satisfaction is the highest, followed by agreement, sexuality, and independence. Correlations between the 4 subscales and relationship satisfaction in this study depict however an interesting deviation from the reported pattern (Table 2). All in all, in accordance with the above-mentioned results, correlations within the LFR group are lower than those within the HFR group. More important, however, is the difference with respect to intimacy. In the LFR group, intimacy and relationship satisfaction correlated significantly lower (r = .61, p < .01) than in the HFR group (*r* = .89, *p* < .001), *z* = 2.02, *p* < .05). Emotional support, trust, and understanding, to name a few, are significantly less important in the low fertile phases of the female cycle.

It is not surprising that the correlation between sexuality and relationship satisfaction is greater in the HFR group than in the LFR group. However, this difference is not significant, z = 1.41, p > .14. The pattern of results for independence was not expected. Women in the fertile phases based relationship satisfaction evaluation more on independence than women in the low fertile phases.

Discussion

The maintenance of a satisfactory relationship is a task that is distinguished by high personal relevance for both men and women, and a wide variety of emotional and social consequences are associated with a stable, satisfactory relationship. Relationships however present different tasks for men and women—not only against an evolutionary background. In societies where women have poorer access to educational facilities and socio-economical resources than men, a long-term romantic relationship seems to offer women the opportunity of upward mobility, in which women exchange one of their few possessions, their "capacity to give birth", for material security.

The fact that women value different characteristics than men when selecting their partner is comprehensible, and is well documented in the literature (for example, BOTWIN/BUSS/SHACKELFORD 1997; TOWNSEND/WASSERMAN 1998; KASSER/SHARMA 1999, to name a few newly conducted studies; for a gen-

Evolution and Cognition | 119 | 2003, Vol. 9, No. 2

Explained variance in high fertility group higher than in low fertility group				
	Explained Variance (r²)	Explained Variance (r²)	ference	
Relation Quality Feature	LFG	HFG	Dif	
Tenderness	.14	.85	.71	
Being there for each other	.19	.88	.69	
Security	.09	.78	.69	
Forgiveness	.00	.69	.69	
Understanding	.17	.78	.61	
Sexual satisfaction	.07	.68	.61	
Willingness to compromise	.12	./2	.60	
Accepting partner	.01	.60	.59	
Longing for each other	.17	./3	.30	
Elevibility	.11	.03	.32	
Paying attention to partner	.00	.40	.40	
Discussing everything	.15	.07	47	
Empathy	25	.03	46	
Dependability	.01	.46	.45	
Friendship	.04	.49	.45	
Sexuality	.29	.71	.42	
Taking interest in partner	.45	.84	.39	
Honesty	.00	.37	.37	
Harmony	.38	.68	.30	
Taking time for each other	.51	.81	.30	
Safety	.52	.81	.29	
Physical contact	.45	.73	.28	
Support	.27	.53	.26	
Willingness to discuss things	.10	.35	.25	
Displaying emotions	.18	.42	.24	
Affection	.61	.84	.23	
Deferring to partner' wishes	.03	.26	.23	
Mutual respect	.47	.69	.22	
Trust	.55	.75	.20	
Humor	.05	.25	.20	
Consideration	.02	.21	.19	
Equality	.40	.59	.19	
Spending as much time together as possible	.20	.37	.17	
Talking with each other	.23	.40	.17	
Mutual friends	.49	.64	.15	
Maintaining individuality	.00	.14	.14	
Helping one another	.44	.58	.14	
Tolerance	.26	.39	.13	
Sexual harmony	.62	.75	.13	
Willingness to argue when necessary	.05	.16	.11	
Love	.69	.78	.09	
Common activities	.32	.41	.09	
Solving problems together	.53	.62	.09	
Arguments	.00	.07	.07	
Having run	.27	.33	.06	
Responsibility	.02 40	.07	.03	
Autonomy	.49	.33	.04	
Openness	.01	.03	.04	
Openness Having and allowing for freedom	.32 00	.33	.03 02	
Own friends	.00	.02	.02	
Own menus	.07	.00	.01	

Explained variance in low fertility group higher than in high fertility group					
Relation Quality Feature		Explained Variance (r²)	fference		
Relation Quality Feature	LFG	HFG	Di		
No dominance	.04	.02	02		
Fidelity	.34	.31	03		
Not taking each other for granted	.07	.03	04		
Only a few quarrels	.10	.06	04		
Looking forward to seeing each other	.55	.49	06		
Listening to each other	.27	.15	12		
Similar beliefs	.55	.42	13		
Different interests	.18	.02	16		
Similar interests	.21	.02	19		
Mutual goals	.43	.13	30		
Running the household together	.50	.04	46		
Similarities	.54	.00	54		

Table 1. Relationship quality features explain differentamounts of relationship satisfaction variance in high fertility(HFG) and low fertility (LFG) women.

	LFG (<i>N</i> = 19)	HFG (<i>N</i> = 20)
Intimacy	.61	.89
Agreement	.85	.88
Sexuality	.77	.87
Independence	.73	.89

Table 2. Correlations between dimensions of relationship quality scales and relationship satisfaction for low and high fertility risk groups. Note: all p < .01, two-tailed.

eral review see FEINGOLD 1992). There is some controversy in the literature as to what determines these differences (for example BUSS 1989 vs. KASSER/ SHARMA 1999).

These sex differences may easily be explained with differences in minimal parental investment. They could also well be embedded within a sociocultural perspective (e.g., HOWARD/BLUMSTEIN/ SCHWARTZ 1987). In cultures where women are more financially dependent on men-rather than viceversa-they are likely to have more to lose when their relationship comes to an end. Given this situation, women are likely to adopt a pragmatic and realistic point of view when it comes to evaluating their relationship. The fact that they cannot afford "the luxury of being romantic" is hence quite easy to grasp, and is consistent with the observation that women seem to be better "barometers of their relationship" (BENTLER/NEWCOMB 1978; RUVOLO/VEROFF 1997; Sprecher 2001).

Evolution and Cognition | **120** | 2003, Vol. 9, No. 2

In this study, I found that relationship scrutiny was linked to the menstrual cycle of the women. Differences between high and low fertility risk women were in fact obtained (a) in the correlations of a relationship quality composite and relationship satis-

faction, (b) in the number of single order correlations between relationship quality features and relationship satisfaction, and finally in correlations between four dimensions of relationship quality and relationship satisfaction. All of these methods converge to the same result: High fertility risk women scrutinized their relationship more than low fertility risk women did.

These differences are in fact more remarkable given the fact that my method in dividing women into high fertility and low fertility risk groups may have been afflicted with inaccuracy. This may on the one hand be inaccurate as the fertility phase between the 6th and 14th day of the menstruation cycle is a widely defined range, and on the other hand since the women had to state the number of days since their last menstrual cycle in a retrospective manner. These inaccuracies may in fact have undermined the actual differences between high and low fertility risk groups. The inaccuracies could not however be used as a counter argument to the reported findings, as a more reliable assessment of fertility would have resulted in less error of measurement. One would expect a far more transparent picture to emerge, precisely elucidating relationship-relevant information processing along the menstrual cycle by women, if for example, the precise day of ovulation was determined with the use of hormone-based tests.

In my opinion, the influence of fertility risk on the processing and cognitive integration of relationship-relevant information can be parsimoniously explained within an evolutionary framework. From an evolutionary perspective, these differences between low fertility risk and high fertility risk women make sense. When, for example, the probability of conception is high, a woman carefully examines the potential father of her child, scrutinizes her rela-

Author's address

Manfred Hassebrauck, Fachbereich G, Sozialpsychologie, Bergische Universität Wuppertal, Gaussstrasse 20, D-42097 Wuppertal, Germany. Email: hassebrauck@uni-wuppertal.de tionship, and questions whether the features necessary to raise a child are truly available.

Although evolutionary and sociocultural perspectives are not opposing views, the effects of the menstrual cycle cannot be accounted for

here by a sociocultural perspective. On a day during their high fertile phase, women process information concerning their relationships systematically and carefully, and regard sexuality and similarities, just to mention a few features, as important factors of a satisfactory relationship. Another group of women, during the low fertility phases of their cycle, processes relationship information less systematically and less carefully and places more weight on features such as similarity and less value on sexuality.

The fact that women are more willing to cheat on their partner (GANGESTAD/THORNHILL/GARVER 2002) during their fertile period does not contradict the present findings. On the contrary—extra-pair mating involves the risk of losing the primary mate. Against this background, it is only adaptive to analyze one's relationship thoroughly *before* they endanger it. It would be of interest for further research to examine whether women who cheat on their partner during their fertile days—or could at least imagine to do so—had analyzed their relationship thoroughly and as a result established that it is in contrast to their ideal type of relationship.

Taking women's fertility status into account in research in close relationships may shed some light on otherwise unknown and unnoticed facts. This variable could possibly explain the many other inconsistent findings in the literature on sex differences research.

Acknowledgment

This research was supported by a grant from the German Science Foundation (Deutsche Forschungsgemeinschaft). I would like to thank Jennifer AUSTIN and Helen GREENLAND for their assistance with the translation.

References

- Allgeier, E. R./Wiederman, M. W. (1994) How useful is evolutionary psychology for understanding contemporary human sexual behavior? Annual Review of Sex Research 5: 218–256.
- Baker, R. R./Bellis, M. A. (1995) Human Sperm Competition: Copulation, Masturbation, and Infidelity. Chapman & Hall: London.
- Bentler, P./Newcomb, M. (1978) Longitudinal study of marital success and failure. Journal of Consulting and Clinical Psychology 46 1053–1070.
- Botwin, M. D./Buss, D. M./Shackelford, T. K. (1997) Personality and mate preferences: Five factors in mate selection and marital satisfaction. Journal of Personality 65: 107–136.
- Buss, D. M. (1989) Sex differences in human mate preferences: Evolutionary hypotheses tested in 37 cultures. Behavioral and Brain Sciences 12: 1–49.
- Buss, D. M. (1994) The Evolution of Desire: Strategies of Human Mating. Basic Books: New York.
- **Buss, D. M./Schmitt, D. P. (1993)** Sexual strategies theory: An evolutionary perspective on human mating. Psychological Review 100: 204–232.
- Campbell, W. K./Sedikides, C./Bosson, J. (1994) Romantic involvement, self-discrepancy, and psychological well-being: A preliminary investigation. Personal Relationships 1: 399–404.
- Clore, G. L./Schwarz, N./Conway, M. (1994) Affective causes and consequences of social information processing. In: Wyer, R. S./Srull, T. K. (eds) Handbook of Social Cognition (2nd edition). Erlbaum: Hillsdale NJ, pp. 323–417.
- **Ellis, B. J. (1992)** The evolution of sexual attraction: Evaluative mechanisms in women. In: Barkow, J./Cosmides, L./ Tooby, J. (eds) The Adapted Mind. Oxford University Press: New York, pp. 267–288.
- Feingold, A. (1992) Gender differences in mate selection preferences: A test of the parental investment model. Psychological Bulletin 112: 125–139.
- Feldman, S. S./Fisher, L./Seitel, L. (1997) The effect of parents' marital satisfaction on young adults' adaptation: A longitudinal study. Journal of Research on Adolescence 7: 55–80.
- Gangestad, S. W./Cousins, A. J. (2002) Adaptive design, female mate preferences, and shifts across the menstrual cycle. Annual Review of Sex Research 12: 145–185.
- Gangestad, S. W./Thornhill, R./Garver, C. E. (2002) Changes in women's sexual interests and their partners' mate-retention tactics across the menstrual cycle: Evidence for shifting conflicts of interest. Proceedings of the Royal Society of London B 269: 975–982.
- Hassebrauck, M. (1991) ZIP—Ein Verfahren zur Erfassung der Zufriedenheit in Paarbeziehungen. Zeitschrift für Sozialpsychologie 22: 256–259.
- Hassebrauck, M. (1997) Cognitions of relationship quality: A prototype analysis of their structure and consequences. Personal Relationships 4: 163–185.
- Hassebrauck, M./Aron, A. (2001) Prototype matching in close relationships. Personality and Social Psychology Bulletin 27: 1111–1122.
- Hassebrauck, M./Fehr, B. (2002) Dimensions of relationship quality. Personal Relationships 9: 253–270.
- Hassebrauck, M./Sedikides, C. (under review) The scrutiny model of relationship evaluation.
- Hatfield, E./Sprecher, S. (1995) Men's and women's preferences in marital partners in the United States, Russia, and Japan. Journal of Cross-Cultural Psychology 26: 728–750.
- Hendrick, S. S. (1988) A generic measure of relationship satis-

faction. Journal of Marriage and the Family 50: 93–98.

- Hobart, C. W. (1958) Incidence of romanticism during courtship. Social Forces 36: 363–367.
- Howard, J. A./Blumstein, P./Schwartz, P. (1987) Social or evolutionary theories? Some observations on preferences in human mate selection. Journal of Personality and Social Psychology 53: 194–200.
- **Kasser, T./Sharma, Y. S. (1999)** Reproductive freedom, educational equality, and females' preference for resource-acquisition characteristics in mates. Psychological Science 10: 374–377.
- **Kephart, W. (1967)** Some correlates of romantic love. Journal of Marriage and the Family 29: 470–479.
- Kiecolt-Glaser, J. K./Fisher, L. D./Ogrocki, P. (1987) Marital quality, marital disruption, and immune function. Psychosomatic Medicine 49: 13–34.
- **Kurdek, L. A. (1993)** Predicting marital dissolution: A 5-year prospective longitudinal study of newlywed couples. Journal of Personality and Social Psychology 64: 221–242.
- Kurdek, L. A. (2002) Predicting the timing of separation and marital satisfaction: An eight-year prospective longitudinal study. Journal of Marriage and the Family 64: 163–179.
- Mayne, T. J/ OLeary, A./McCrady, B./Contrada, R./Labouvie, E. (1997) The differential effects of acute marital distress on emotional, physiological and immune functions in maritally distressed men and women. Psychology & Health 12: 277–288.
- **Penton-Voak, I. S./Perrett, D. I. (2000)** Female preference for male faces changes cyclically: Further evidence. Evolution and Human Behavior 21: 39–48.
- **Regan, P. C. (1996)** Rythms of desire: The association between menstrual cycle phases and female sexual desire. Canadian Journal of Human Sexuality 5: 145–156.
- Russell, R. J. H./Wells, P. A. (1994) Predictors of happiness in married couples. Personality and individual differences 17: 313–321.
- **Ruvolo, A. P./Veroff, J. (1997)** For better or for worse: Real-ideal discrepancies and the marital well-being of newlyweds. Journal of Social and Personal Relationships 14: 223–242.
- Schwarz, N./Bless, H. (1991) Happy and mindless, but sad and smart? The impact of affective states on analytic reasoning. In: Forgas, J. P. (ed) Emotion and Social Judgments. Pergamon Press: Oxford, pp. 55–71.
- Sprecher, S. (2001) A comparison of emotional consequences of and changes in equity over time using global and domain-specific measures of equity. Journal of Social and Personal Relationships 18: 477–501.
- Sprecher, S./Metts, S. (1989) Development of the 'romantic beliefs scale' and examination of the effects of gender and gender-role orientation. Journal of Social and Personal Relationships 6: 387–411.
- Symons, D. (1979) The Evolution of Human Sexuality. Oxford University Press: New York.
- **Thornhill, R./Gangestad, S. W. (1999)** The scent of symmetry: A human sex pheromone that signals fitness? Evolution and Human Behavior 20: 175–201.
- Townsend, J. M./Wasserman, T. (1998) Sexual attractiveness: Sex differences in assessment and criteria. Evolution and Human Behavior 19: 171–191.
- Trivers, R. L. (1972) Parental investment and sexual selection. In: Campbell, B. (ed) Sexual Selection and the Descent of Man: 1871–1971. Aldine: Chicago, pp. 136–179.
- Wickrama, K. A. S./Lorenz, F. O./Conger, R. D./Elder, G. H. (1997) Marital quality and physical illness: A latent growth curve analysis. Journal of Marriage and the Family 59: 143–155.

Evolution and Cognition | **122** | 2003, Vol. 9, No. 2

Darwin and Descartes' Demon

On the Possible Evolutionary Origin of Belief in an External World

Without any reasoning, or even almost before the use of reason, we always suppose an external universe which depends not on our perception but would exist though we and every sensible creature were absent or annihilated. (HUME 1955, Section XII, Part I)

Those cousins of our ancestors who could not manage to learn that there was an independently existing 'external world,' one whose objects continued on trajectories or in place even when unobserved, did not fare as well as those who quickly recognized obdurate realities. (NOZICK 1993, p121)

How can we be certain that our senses provide us with an accurate picture of the world? How can we be certain that there really is an objective world at all, and that all our exAbstract

This article explores the possibility that the tendency to believe in an objective, mind-independent external world traces to innate aspects of the human mind. The aspects of mind in question are, first, the capacity to distinguish mental states that have objective referents from those that do not (e.g., perceptual states versus mental imagery), and second, the capacity to mentally represent the continued existence of parts of the world that are beyond the reach of the senses. It is proposed that the evolutionary function of these cognitive abilities relates to the production of novel but adaptive voluntary behaviour. Evidence and arguments are provided in support of the innateness hypothesis. Among these is a Chomskyan-style poverty-of-thestimulus argument derived from the philosophical literature. The evolutionary account of the subjective-objective distinction leads to the prediction that, in conditions of uncertainty, people will tend to err on the side of assuming the objectivity of their perceptions and other judgements.

Key words

Evolutionary psychology, external world, innate ideas, object permanence, objective-subjective distinction, philosophy.

perience is not simply a vivid dream or hallucination? These questions relate to one of the most fundamental and long-standing issues in philosophy: the problem of justifying belief in an external world. The goal of this article is not to resolve this difficult philosophical conundrum. My focus is psychological rather than philosophical, and I will simply assume what many philosophers believe is impossible to prove: that there is an external world, causing our perceptions and existing independently of them. In making this assumption so casually, I am engaging in precisely the phenomenon that is the focus of the article. My contention is that the most intuitively plausible position on this issue for the vast majority of humankind is metaphysical realism, the view that there is an independently existing external world. A number of philosophers have suggested that the disposition to form this view is a part of human nature (Hume 1978; NOZICK 1993; PUTNAM 1981). In this article, I provide a preliminary and highly speculative sketch of how these suggestions might ultimately be placed within an evolutionary psychological framework. I explore the possibility that the assumption of an external

world relates to certain innate aspects of the human mind (that is, aspects that were not derived solely from experience, but can be traced, at least in part, to information contained in materials of the developmental process, in particular the genes; for further discussion of innateness, see STEWART-WILLIAMS submitted). My goal is not to establish this position beyond reasonable doubt. I will be content if I can show simply that the position is not unreasonable, and provide the groundwork for further discussion.

Evolution and Cognition | **123** | 2003, Vol. 9, No. 2

The Evolutionary Function of Common Sense Realism

Metaphysical realism incorporates two closely related assumptions. The first concerns what is perceived, the second what is not perceived. (1) Objectivity. The world revealed by the senses is real; we perceive objective parts of reality. (2) Mind-independence. Not only are the things we perceive real; reality continues to exist when unperceived. (So, for instance, SCHRÖDINGER's cat is either dead or alive before anyone opens the box.) My position is not that these abstract verbal formulations are directly innate. What I propose is that each of our common sense assumptions-objectivity and mind-independence-can be traced to an aspect of mind that evolved to solve distinct adaptive challenges faced by our hominid or pre-hominid ancestors. The aspects of mind I propose are as follows. First, I suggest that human beings have a sense or feeling of the reality of their perceptual experiences, and that this reflects the operation of an evolved mechanism designed to 'label' activity of the sensory regions of the brain as either objective or subjective. Second, I suggest that human beings possess evolved psychological mechanisms designed to represent the continued existence of objective aspects of reality (for instance, physical objects) when these are beyond the range of the senses. In the following sections, I flesh out these speculations, and provide an evolutionary rationale for each.

The Objective–Subjective Distinction

To begin with, I will examine the idea that we possess a perceptual-level sense of the reality of our sensory experience, and that this is part of an evolved mechanism for distinguishing the objective from the subjective. Whereas it is unlikely to be a controversial claim that people assume the continued existence of unperceived aspects of the world, it might be questioned whether we do in fact possess a sense that some of our perceptions are veridical. The argument could be made that, pre-philosophically, we simply act *as if* we hold this belief, but do not actually hold it at all. My first task, then, is to demonstrate that the assumption of objectivity is not merely implicit in our behaviour, but can be traced to genuine mental content.

Support for this conclusion comes from a consideration of atypical states of consciousness. Deviations from the norm can sometimes reveal things about the mind that we might not previously have

been aware of. Consider the capacity to construe people as possessing thoughts, feelings, and intentions (theory of mind). It is thought that most human beings have this capacity (STEWART-WILLIAMS in press). However, because it comes so naturally to us, we may tend to overlook this fact until we encounter people in whom it is absent, such as people with severe autism (BARON-COHEN 1995). The general lesson is that atypical states or conditions may awaken us to aspects of our minds that are so ubiquitous in our experience that we tend not to notice them. An atypical psychological state that may be relevant to the issue of our construal of the external world is the dissociative state of *derealization*. This experience occurs in a number of clinical conditions, and can also occur when people face novel or unusual circumstances, or experience grief or other strong emotions. Individuals experiencing derealization have the sense that the external world is not really real, and often feel that they are in a dream (DAVI-SON/NEALE 2001; KAPLAN/SADOCK 1998). Derealization may reveal something about typical human experience that we might not have noticed: that usually we do have a sense of the reality of the world. The world around us seems real in a way that memories and sensory imagery do not. This argues against the view that the assumption of the objectivity of perception is merely implicit in species-typical responses to perceptual experience. Derealization indicates that we possess genuine mental content related to the objectivity of our perceptual experience, even if we have not reflected on this issue.

The next step is to argue that this aspect of human phenomenology reflects the operation of an evolved psychological mechanism designed to deal with the adaptive challenge of distinguishing aspects of brain activity that relate to external circumstances (i.e., veridical perceptions) from those that do not (e.g., memories and other mental imagery). There is evidence that the same parts of the brain that become more active in response to an external stimulus also become more active when one simply imagines that same stimulus (ISHAI 1997; KOSSLYN 1994). This raises the possibility that the one could be mistaken for the other, and that, in some sense, the perceptual system must tag sensory experience either as 'real/objective' or as 'subjective'. If this tendency has an evolutionary origin, it must exert some influence on voluntary behaviour. Aspects of mind that do not influence behaviour would have no implications for an organism's inclusive fitness, and could not evolve (BUSS 1999). As such, the first

Evolution and Cognition | 124 | 2003, Vol. 9, No. 2

task in providing a DARWINIAN account of the objective–subjective distinction is to specify how the capacity to make this distinction might link to adaptive voluntary behaviour.

In brief, my suggestion is this: Real and imagined stimuli typically call for different behavioural responses. Veridical perceptions often require an immediate response, whereas sensory imagery does not. As a result, it is necessary for an agent to discriminate objective from subjective states (or rather, subjective states that correspond to objective circumstances from subjective states that do not). People unable to differentiate between the subjective and the objective with a reasonable degree of reliability would either fail to respond to many important elements of their environments, or inappropriately respond to imagery and explicit memories as if it were real. This would clearly put them at a selective disadvantage. Individuals capable of making the distinction reliably would typically have survived, reproduced, and aided kin more successfully than those who were not, and thus the genes underlying this capacity would have increased in frequency relative to alleles.

Admittedly, veridical perceptions are usually more detailed and complex than mental imagery, and it might be argued that there is little danger that the one could be mistaken for the other. However, the differences between perception and imagery do not mean that the subjective-objective distinction is unnecessary; they just mean that the distinction may usually be easy to make. Indeed, the level of detail and complexity of our experience may be one means by which the perceptual system 'decides' whether a given aspect of brain activity should be classed as objective or as subjective. Derealization may involve the temporary malfunctioning of the mechanism involved in making this decision: The brain may misclassify genuine perceptual material as mere mental imagery. Conversely, the auditory hallucinations commonly found in individuals suffering schizophrenia may be a product of the misclassification of one's own inner speech as an objective aspect of the external world (HADDOCK/TARRIER/ SPAULDING 1998; see FRITH/DOLAN 1996, for an account of the neurological correlates of this clinical deficit).

So, the evolutionary function of the objectivesubjective distinction may relate to the fact that perceptual states require different behavioural responses than mental imagery. This suggestion takes the use of mental imagery in human thought as a given. Clearly, though, this is also something that must be explained, and a DARWINIAN explanation may be appropriate. Without going into too much detail about this matter, which is beyond the scope of the present exploration, it is plausible that the ability to use explicit memories and mental imagery in thinking evolved as a result of its contribution to long-term behavioural planning. Whatever the reason, though, the link between this ability and the subjective-objective distinction has several implications. First, it allows us to look more deeply into the evolutionary origins of the capacity to distinguish between the objective and the subjective. It can be presumed that the capacity for image-based thought must have evolved hand-in-hand with the capacity to distinguish veridical perception from mere imagery. Therefore, the subjective-objective distinction may ultimately have been selected because it made possible mental imagery and explicit memory. Second, it seems unlikely that most other animals use mental imagery or explicit memories in their behavioural planning, and thus we would not expect them to make the objective-subjective distinction. It may be difficult for us to imagine, but it is possible that most other animals *perceive* their environment, but do not 'label' their perceptual states as objective. This is not to say that they experience them as subjective; instead, they may simply not make the distinction.

It is important to stress that I am not arguing for the ultimate validity (or otherwise) of the objective– subjective distinction, or that every aspect of mental life can be placed neatly and unambiguously into one or the other category. My suggestion is simply that the distinction may have an evolutionary origin. This position only requires that the distinction is valid enough to be useful, and that in evolutionarily relevant cases, people will usually class a given aspect of mental experience in the same way. For example, when there really is a tiger present, most people in most cultures will class the resulting perceptual experience as veridical and objective, rather than as a memory or mental image.

Mind-Independence

Another reason that it is necessary to differentiate between the objective and the subjective is this: It is often adaptively useful to track the continued existence of objective aspects of the world that are beyond the range of the senses, whereas this does not apply in the case of sensory imagery. The ability to track aspects of the environment is widespread in the biological world.¹ DENNETT (1991) goes so far as

Evolution and Cognition | **125** | 2003, Vol. 9, No. 2

to compare our tracking abilities to the simpler occurrence of a sunflower tracking the movement of the sun across the sky. He recognizes, of course, that there are important differences: "If the sun is temporarily obscured, the sunflower cannot project the trajectory; the mechanism that is sensitive to the sun's passage does not represent the sun's passage in this extended sense" (p191). In contrast, we are able to represent items even when we are not in direct causal contact with them, and are able to track their probable paths. Without this ability, we would in effect be solipsists, treating the immediate evidence of the senses as if it were all that is real. The ability to represent items that are not immediately present to us is thus an important element of everyday realism.

It is not difficult to see how the understanding that unperceived parts of the world continue to exist would be adaptive. To provide a simple illustration, a predator without this understanding would stop pursuing its prey the instant it slipped from view. Solipsistic prey would be similarly disadvantaged. Looking more deeply into the issue, SOBER (1994) makes an interesting suggestion concerning the adaptively relevant difference between realism and solipsism. His contention is that solipsists can only make inductive inferences based on regularities in their own past experience, whereas non-solipsists can also make abductive inferences (inferences to the best explanation) for these regularities. Abductive inference requires the understanding that reality extends beyond what can be perceived. According to SOBER (1994), one of the evolutionary advantages of abduction is that, whereas induction simply allows us to generalize from past experiences to similar experiences, abduction allows us to make sense of novel experiences. So, for instance, a realist who came across a dead antelope for the first time could infer that a tiger might be nearby, and act accordingly. To a solipsist in the same situation, however, no such inference would be available. The solipsist would clearly be at a selective disadvantage.

The consciously accessible understanding of the mind-independence of objective reality may underpin one of the most characteristic aspects of human behaviour: our behavioural flexibility. SOBER (1994) suggests that "selection will favor nonsolipsism over solipsism only when innate structures are an insufficient guide to effective behavior" (p38). Thus, if we were innately 'hardwired' to run away both from tigers and from dead antelopes, there would be no further evolutionary advantage in understanding that the dead antelope indicates that a tiger might be nearby. However, because humans have

few innate behaviours and a high level of behavioural flexibility and novelty, a more abstract and general understanding of the world is necessary an understanding that can underlie a wide variety of novel behaviours. Although these behaviours themselves are not innate, the capacity to understand that the world goes on without us may be. This understanding is of course part of our conscious picture of the world. Conscious aspects of the brain's activity tend to be associated with the execution of novel, unpractised behaviours. This is consistent with the view that this understanding is an ingredient in the formulation of novel behaviours.

Evidence and Arguments For and Against Innateness

If nothing else, then, the assumption of an objective, mind-independent external world makes good evolutionary sense. In this section, I will explore evidence and arguments for and against the view that this feature of mental life does indeed have an innate origin. Evolutionary psychologists typically assume that complex products of selection are speciestypical (TOOBY/COSMIDES 1992). As such, to make the case that belief in an external world traces to innate aspects of the mind, I must first provide reason to believe that most people do in fact hold this belief. Universality is not sufficient to infer innateness, but it is commonly viewed as a necessary precondition (GEARY 1995; PINKER 1994).² First, I will deal with some potential criticisms of the idea that belief in an external world is universal. One is the idea that this belief is unique to modern Western thought. Various thinkers have hinted at this view, the truth of which would undermine the innateness hypothesis. For instance, VAN INWAGEN (1993) has described belief in a mind-independent external world as an element of what he calls the 'Common Western Metaphysic'. The implication is that this view is a part of modern Western culture, and not necessarily found beyond this context. However, there is some reason to question this supposition. For a start, the view is not unique to modern Western thought. As NOZICK (2001) notes, the idea that "the world exists in a definite state independently of our observations" (p133) was traditionally viewed as so certain that it was classed as a metaphysical necessity.3 Furthermore, indirect evidence suggests that these intuitions are also found in non-Western cultures. In all languages that have been examined, a distinction is drawn between dreaming and reality (BROWN 1991). Without an underlying belief in an

Evolution and Cognition | 126 | 2003, Vol. 9, No. 2

external world, it is not clear what this distinction would amount to.

But even if some people in every culture hold this belief, it might be argued that there are surely many who do not. There are some philosophers, for instance, who doubt or deny the reality of an external world (sceptics), and others who deny that the world is mind-independent (anti-realists). This appears to be inconsistent with the innateness hypothesis. However, again there are reasons not to give this argument too much weight. HUME (1978) pointed out that, although sceptical philosophers claim not to believe in an external world, in their everyday reasoning and actions they act suspiciously like they do. A similar point was made by Bertrand RUSSELL (1927), who was fond of telling the story of the woman who claimed to be a solipsist, but wondered why more people were not also. Postmodernists are a contemporary example. Like traditional idealists or solipsists, some postmodernists flatly deny the reality of the external world, but in giving lectures and preparing journal articles on this theme, their actions contradict their stated beliefs (GOLDMAN 1999). To borrow a phrase used by HUME (1978) in another context, it appears that the belief in an external world is "one of those maxims, which tho' they may be deny'd with the lips, 'tis impossible for men in their hearts really to doubt of" (Book I, Part III, Section III, Paragraph 1). Furthermore, an evolutionary perspective puts an interesting new spin on the issue of radical scepticism. Consider HUME's comments about his own sceptical inquiries:

"Most fortunately it happens, that since reason is incapable of dispelling these clouds, nature herself suffices to that purpose, and cures me of this philosophical melancholy and delirium, either by relaxing this bent of mind, or by some avocation, and lively impression of my senses, which obliterate all these chimeras. I dine, I play a game of back-gammon, I converse, and am merry with my friends; and when after three or four hour's amusement, I wou'd return to these speculations, they appear so cold, and strain'd, and ridiculous, that I cannot find in my heart to enter into them any farther." (HUME, 1978, Book I, Part IV, Section, VII, Paragraph 9)

The experience HUME describes might profitably be compared to the phenomenon of 'instinctive drift' (BRELAND/BRELAND 1961, 1966). Early behaviourists argued that the process of reinforcement and punishment could shape essentially any behaviour (SKINNER 1938). Instinctive drift was one among a number of discoveries that challenged this position. It occurs when animals are trained to act in ways that clash with the behaviours typical of their species. In time, the trained behaviour deteriorates, and the animals revert to more natural patterns of behaviour. HUME's famous passage raises the possibility that something similar happens to people who pry their minds away from the natural intuition that there is an external world. We might maintain such a belief momentarily, but in time we may slide back to more 'instinctive' ways of thinking.

These arguments do not establish beyond any reasonable doubt that the belief in an external world is universal across human cultures; they are suggestive rather than conclusive. However, they do call into question some of the main objections that may be brought against this position. As such, the universality hypothesis remains a live option. Although some profess to deny the existence of an objective world, it is reasonable to think that metaphysical realism is the intuitive position for most people in most cultures, and that it is a candidate for a species-typical component of our worldview. As mentioned, though, the universality of this position is not adequate evidence for its innateness. So the next question is: What is the origin of our realist bent? Does it derive from experience, or is there an innate contribution?

There is unfortunately little evidence bearing on the issue of the innateness or otherwise of the subjective-objective distinction. At this stage, the best reason to entertain the innateness hypothesis is simply that it makes good evolutionary sense. However, a stronger case can be made in support of the innate origin of the understanding that the world continues to exist when unperceived. My main argument for this conclusion is a poverty-of-the-stimulus argument. This draws on epistemological arguments in the philosophical literature, but turns them to a new purpose. Philosophers point out that no evidence could prove beyond a shadow of doubt that our sensory experience is not all simply a dream or hallucination, or that the world continues to exist when we close our eyes, only to spring back into existence when we open them again. Such possibilities are usually raised in the context of asking how we can justify our belief that sensory perception is veridical. But they also raise another question, one more psychological than philosophical: If there is no evidence for this belief, how do we form the belief in the first place? Consider the assumption of object permanence. As BERKELEY (1982) and others have noted, it is simply not possible to perceive an unperceived object. Nonetheless, with the

Evolution and Cognition | **127** | 2003, Vol. 9, No. 2

(possible) exception of sceptical philosophers, we all assume that objects continue to exist when beyond the range of our senses. Logically, solipsism is as valid an interpretation of our sensory experience as the view that there are objects existing independently of perception. Given that our experience is consistent with solipsism, the assumption of object permanence could not be derived solely from experience. This suggests that there must be an innate contribution.

It might be argued, though, that the assumption that objects persist is a more useful interpretation of the evidence than the solipsist's interpretation, and that this is why people favour realism over solipsism. Without denying that it is a more useful interpretation, it is not clear that this proposal could be correct. A working understanding of object permanence has been established in a variety of nonhuman animals (WYNNE 2001). Furthermore, research suggests that as soon as infants are old enough to be tested (around three months), they represent the continued existence of objects beyond their field of view (BAILLARGEON 1999; BAILLARGEON/SPELKE/ WASSERMAN 1985; CAREY/SPELKE 1994; SPELKE 1990; although see HAITH 1998).4 But it does not seem reasonable to suggest that nonhuman animals or human infants reason their way to the most useful interpretation of the evidence. These considerations lend some support to the notion that the assumption of object permanence-one of the key elements in the understanding of the mind-independence of objective reality-has an innate origin.

Assuming Objectivity Over Subjectivity: An Adaptive Bias?

According to evolutionary psychologists, humans have evolved a number of adaptive cognitive biases. For instance, in the identification of threats, people (and other animals) tend to err on the side of making false positives (LEDOUX 1996). Similarly, people err on the side of making animistic or anthropomorphic interpretations of ambiguous stimuli (GUTHRIE 1994). The evolutionary rationale for these biases is that it is less costly to make false positives than to make false negatives-for instance, it is less costly to judge that a tiger is present when it is not than to judge that it is not present when it is—and that selection will favour a tendency to err on the side of the less costly error (HASELTON/BUSS 2000). If the tendency to tag aspects of experience as objective or subjective has an evolutionary origin, it would be predicted that we possess another adaptive bias:

Where there is any doubt or ambiguity, we will assume the objectivity of our perceptual experience. Admittedly, mistaking mere imagery for veridical perception would not be a fitness-enhancing strategy. However, mistaking perception for imagery would probably be an even greater threat to the inclusive fitness of our hunter-gatherer ancestors. WISNIEWSKI (1998) puts the point simply: "a person contemplating the veracity of their percepts might be eaten by an approaching tiger" (p57). This leads to the prediction that, under conditions of uncertainty, we will tend to err on the side of assuming that our sensory-perceptual experiences are veridical. The same argument applies to other judgements about the world. RUSE (1986) sums up the position: "A tendency to objectify is the price of reproductive success" (p172).

Preliminary evidence for the existence of this bias can be found in your own experience: Which has more common for you, mistaking a dream for reality or mistaking reality for a dream? In support of the view that the objectivity bias extends beyond perceptual states to other judgements, the anthropologist Donald BROWN (1991) reports that people in all cultures overestimate the objectivity of thought. One example is that we tend to view matters of taste as having objective validity. KANT (1952) noted that aesthetic subjectivism (the idea that beauty is solely in the eye of the beholder) clashes with common sense. GARDNER (1995) elaborates this point:

"Common sense is unequivocal that 'the music is beautiful' means more than, even if it presupposes, 'the music gives me pleasure'. If those two thoughts were the same, one could not regard one's aesthetic judgement as something that another person might take issue with, and in support of which reasons other than the fact of one's liking it—may be given. But we do think that a judgement that the music is beautiful clashes with the judgement that it is ugly or sentimental: it does so because—as its grammatical form suggests—it aims to say something about the music, not about oneself" (p592).

Similarly, people tend to attribute objective validity to their moral judgements, even though no one has yet been able to construct a plausible argument in favour of moral objectivism (JOYCE 2001; MACKIE 1977; NOZICK 1981). Although targeted empirical evidence is needed, these considerations at least render plausible the notion that we possess an adaptive bias toward assuming the objectivity of our perceptions and other judgements. This supports an evolutionary interpretation of the objective–subjec-

Evolution and Cognition | **128** | 2003, Vol. 9, No. 2

tive distinction, which in turn supports the view that the belief in an objective external reality can be traced to innate aspects of mind.

Conclusion

HUME (1955) said: "The mind has never anything present to it but the perceptions, and cannot possibly reach any experience of their connection with other objects" (Section XII, Part I, Paragraph 12). But despite the fact that perception alone reveals nothing about the objectivity or mind independence of the world, the common sense assumption for the majority of people is metaphysical realism. Evolutionary considerations and various lines of evidence raise the possibility that this assumption traces to innate aspects of the human mind, namely, the capacity to distinguish mental events that have objective referents from those that do not, and the capacity to represent the continued existence of unperceived parts of the world. These capacities plausibly enhanced the inclusive fitness of our ancestors. The ability to distinguish the subjective from the objective may relate to the production of appropriate behaviour in response to subjective versus objec-

Notes

1 As noted, it is unlikely that the use of sensory imagery in thought is nearly as common. This has an interesting implication. An ability shared by a number of species may trace to a common ancestor. The more widespread the ability is, the longer ago the common ancestor must have lived and the more ancient the ability. Therefore, assuming that the ability to track objects is more common than the use of sensory imagery in thought, this would argue that the former ability evolved long before the latter.

References

- **Baillargeon, R. (1999)** The object concept revisited: New directions in the investigation of infant's physical knowledge. In: Margolis, E./Laurence, S. (eds) Concepts: Core readings. MIT Press: Cambridge MA, pp. 571–612.
- Baillargeon, R./Spelke, E. S./Wasserman, S. (1985) Object permanence in five-month-old infants. Cognition 20: 191–208.
- Baron-Cohen, S. (1995) Mindblindness: An Essay on Autism and Theory of Mind. MIT Press: Cambridge MA.
- Berkeley, G. (1982) Treatise Concerning the Principles of Hu-

Author's address

Steve Stewart-Williams, School of Psychology, Massey University, Palmerston North, New Zealand. E-mail: anonymous1@xtra.co.nz tive aspects of mental experience, and may be associated with an adaptive bias toward assuming the objectivity of our perceptual and other judgements. The conscious understanding that unperceived parts of the world con-

tinue to exist may relate to tracking environmental regularities, and may be a necessary ingredient in the generation of novel adaptive behaviour. At this stage, these views are highly speculative, and the arguments and evidence provided certainly do not constitute an unassailable proof. Nonetheless, it is my view that they justify further research and discussion on this topic.

Author Note

This paper was supported by a Bright Future Top-Achiever Doctoral Scholarship, administered by FRST. Thanks go to Dr. Stephen HILL and Dr. John PODD (School of Psychology, Massey University), and to James BATTYE (School of History, Philosophy, and Politics, Massey University), for helpful suggestions and stimulating discussion concerning this paper.

- **2** This, at least, is the case for evolutionary products unrelated to reproduction. In the case of evolved psychological mechanisms related to reproduction, there may be important differences between the sexes (BUSS 1999).
- **3** This idea is now challenged by some theoretical interpretations of quantum phenomena, which may help to explain the reputation this area of inquiry has for being counterintuitive.
- **4** This research challenges PIAGET's (1954) early claim that infants prior to two years of age have no conception of a world beyond their subjective experience.

man Knowledge. Hackett: Indianapolis. Originally published in 1710.

- Breland, K./Breland, M. (1961) The misbehavior of organisms. American Psychologist 16: 681–684.
- Breland, K./Breland, M. (1966) Animal Behavior. Macmillan: New York.
- Brown, D. E. (1991) Human Universals. McGraw-Hill: New York.
- **Buss, D. M. (1999)** Evolutionary Psychology: The New Science of the Mind. Allyn & Bacon: Needham Heights MA.
- Carey, S./Spelke, E. (1994) Domain-specific knowledge and conceptual change. In: Hirschfeld, L. A./Gelman, S. A.

Evolution and Cognition | **129** | 2003, Vol. 9, No. 2

(eds) Mapping the Mind: Domain Specificity in Cognition and Culture. Cambridge University Press: New York, pp. 169–200.

- Davison, G. C./Neale, J. M. (2001) Abnormal psychology (8th edition). John Wiley & Sons: New York.
- Dennett, D. C. (1991) Consciousness Explained. Little, Brown & Co: New York.
- Frith, C./Dolan, R. (1996) The role of the prefrontal cortex in higher cognitive functions. Cognitive Brain Research 5: 175–181.
- Gardner, S. (1995) Aesthetics. In: Grayling, A. C. (ed) Philosophy: A Guide through the Subject. Oxford University Press: Oxford, pp. 583–627.
- Geary, D. C. (1995) Reflections of evolution and culture in children's cognition: Implications for mathematical development and instruction. American Psychologist 50: 24–37.
- Goldman, A. I. (1999) Knowledge in a Social World. Clarendon Press: Oxford.
- Guthrie, S. E. (1994) Faces in the Couds: A New Theory of Religion. Oxford University Press: New York.
- Haddock, G./Tarrier, N./Spaulding, W. (1998) Individual cognitive-behavior therapy in the treatment of hallucinations and delusions: A review. Clinical Psychology Review 18: 821–838.
- Haith, M. M. (1998) Who put the cog in infant cognition? Is rich interpretation too costly? Infant Behavior and Development 21: 167–179.
- Haselton, M. G./Buss, D. M. (2000) Error management theory: A new perspective on biases in cross-sex mind reading. Journal of Personality and Social Psychology 78: 81–91.
- Hume, D. (1955) An Inquiry Concerning Human Understanding. Bobbs-Merrill: Indianapolis. Originally published in 1748.
- Hume, D. (1978) A Treatise of Human Nature. Oxford University Press: Oxford. Originally published in 1739.
- Ishai, A. (1997) Visual imagery: Effects of short- and longterm memory. Journal of Cognitive Neuroscience 9: 734– 742.
- Joyce, R. (2001) The Myth of Morality. Cambridge University Press: Cambridge.
- Kant, I. (1952) Critique of Judgement. Translated by J. C. Meredith. Clarendon Press: Oxford. German original published in 1790.
- Kaplan, H. I./Sadock, B. J. (1998) Kaplan and Sadock's Synopsis of Psychiatry (8th edition). Williams & Wilkins: Baltimore MD.
- Kosslyn, S. M. (1994) Image and Brain: The Resolution of the Imagery Debate. MIT Press: Cambridge MA.

- **LeDoux, J. E. (1996)** The Emotional Brain: The Mysterious Underpinnings of Emotional Life. Simon & Schuster: New York.
- Mackie, J. L. (1977) Ethics: Inventing Right and Wrong. Penguin: New York.
- **Nozick, R. (1981)** Philosophical Explanations. Clarendon Press: Oxford.
- Nozick, R. (1993) The Nature of Rationality. Princeton University Press: Princeton NJ.
- **Nozick, R. (2001)** Invariances: The Structure of the Objective World. Belknap Press: Cambridge MA.
- Piaget, J. (1955) The Child's Construction of Reality. Routledge & Kegan Paul: London.
- **Pinker, S. (1994)** The Language Instinct: The New Science of Language and Mind. Penguin: London.
- Putnam, H. (1981) Reason, Truth, and History. Cambridge University Press: Cambridge.
- Ruse, M. (1986) Taking Darwin Seriously: A Naturalistic Approach to Philosophy. Blackwell: Oxford.
- Russell, B. (1927) An Outline of Philosophy. Routledge: London.
- Skinner, B. F. (1938) The Behavior of Organisms. Appleton-Century-Croft: New York.
- Sober, E. (1994) From a Biological Point of View: Essays in Evolutionary Philosophy. Cambridge University Press: New York.
- Spelke, E. S. (1990) Principles of object perception. Cognitive Science 14: 29–56.
- Stewart-Williams, S. (in press) Elements of an evolved conceptual framework. In: Fletcher-Flinn, C. M./Haberman, G. M. (eds) Cognition, Language, and Development: Perspectives from New Zealand.
- **Stewart-Williams, S. (submitted)** Innate ideas after Darwin. Manuscript submitted for publication.
- Tooby, J./Cosmides, L. (1992) The psychological foundations of culture. In: Barkow, J. H./Cosmides, L./Tooby, J. (eds) The Adapted Mind: Evolutionary Psychology and the Generation of Culture. Oxford University Press: New York, pp. 19–136.
- van Inwagen, P. (1993) Metaphysics. Oxford University Press: Oxford.
- Wisniewski, E. J. (1998) The psychology of intuition. In: De Paul, M./Ramsey, W. (eds) Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical inquiry. Rowman & Littlefield: Lanham MD, pp.45–58.
- Wynne, C. D. L. (2001) Animal Cognition: The Mental Lives of Animals. Palgrave: Basingstoke UK.

A Test for Comprehension of False Belief in Chimpanzees

Introduction

Theory of mind (or second order intentionality) has come to be the benchmark for human-like advanced social cognition. Children acquire (or develop) this capacity at around the age of 4-5 years in a relatively abrupt fashion (WIMMER/ Perner 1983; Astington 1993). Ever since WIMMER/ PERNER'S (1983) classic study, the false belief task has become the standard assay for this. A number of such tasks (including the Sally-Ann task and the "Smartie" task) have come to be considered as standard in human developmental psychology. Although a verv large number of studies have used false belief tasks to study theory of mind and

Abstract

Four chimpanzees, 11 autistic adults and 41 children aged 3-6 years old were tested on a nonverbal mechanical analogue of a false belief task under two conditions (true belief and false belief). The task was first benchmarked against a conventional false belief task (the Smartie task) using the children and the autistic subjects. The children and the autistic subjects exhibited the classic results on both tasks: all were able to cope with true belief tasks, but only the oldest children were able to cope with false belief tasks. Taken as a whole, the chimpanzees performed significantly better than the autistic subjects and were not significantly worse than 4-5 year children on the mechanical analogue task. However, only one chimpanzee performed at significantly better than chance levels on both versions of the mechanical analogue task. We infer from these results that chimpanzees' mentalising abilities are, at best, equivalent to those of human children around the age at which they just begin to acquire theory of mind.

Key words

Theory of mind, false belief, chimpanzee, autism.

related abilities in human developmental and abnormal psychology, there has to date been only one published study (CALL/TOMASELLO 1999) that has used this approach with nonhuman animals.

It has, however, been suggested that false belief tasks may be too anthropocentric for animals to cope with. Consequently, studies of social cognition in nonhumans have tended, in recent years, to explore competences (in particular, the knowledge states of other individuals, the difference between guessing and knowing, joint attention and role reversal) that may be prerequisites for theory of mind (e.g., in great apes: PREMACK/WOODRUFF 1978; POVINELLI/NELSON/ BOYSEN 1990; POVINELLI/EDDY 1996; POVINELLI et al. 1998; CALL/TOMASELLO 1998; TOMASELLO/CALL/HARE 1998; HARE et al. 2000; in Old World monkeys: SEY-FARTH/CHENEY 1990; POV-INELLI/PARKS/NOVAK 1991, 1992; CHENEY/SEYFARTH/ PALOMBIT 1996; RENDALL/ CHENEY/SEYFARTH 2000). One difficulty with this approach is that none of these studies would meet the challenge laid down by the human developmental psychologists. None of these tasks, for example, explicitly requires an understanding of false belief for its successful completion.

An alternative approach has been the suggestion that anecdotal evidence for tactical deception might fill this gap. However, this approach has ultimately proved to be less compelling than had initially

been supposed. WHITEN/BYRNE (1988; BYRNE/ WHITEN 1991), for example, collated observations of tactical deception in the primate literature, and found a disproportionately high frequency of deception in chimpanzees and baboons. However, not all recorded cases of deceptive behaviour may be second-order intentional, since many could arise through conditioning and associative learning (HEYES 1993, 1998). Evidence based on the frequencies with which animals give food calls, for example, suggests that animals may not always take the audience's state of mind into account: calls may be correlated with the amount of food present rather than with another animal's awareness of the food (HAUSER/WRANGHAM 1987).

Evolution and Cognition | 131 | 2003, Vol. 9, No. 2

Although studies of phenomena like joint attention and intentional behaviour are highly suggestive, most human psychologists accept WIMMER/ PERNER's (1983) argument that such phenomena do not provide unequivocal evidence for theory of mind. Only a false belief task provides that kind of security of inference because false belief tasks can only be solved correctly if the subject can aspire to second order intentionality (i.e., formal theory of mind): the subject has to be able to recognise that another individual has a belief about the world which it supposes to be untrue. Since developmental psychologists have, in effect, set the agenda in this respect, comparative psychologists are obliged to rise to the challenge either by designing nonverbal false belief tasks for animals that meet this criterion or by convincing developmental psychologists that an alternative criterion would be acceptable. Although work on the "seer-knower" paradigm (e.g., HARE et al. 2000, 2003; HARE/CALL/TOMASELLO 2001) is highly suggestive, there seems little serious prospect of the second, so that for the foreseeable future the first is the only realistic alternative if comparison between humans and nonhumans is to have any kind of place on the wider agenda.

In this paper, we report the results of a study using a nonverbal mechanical analogue of a false belief task with chimpanzees. In developing such analogue tasks, it is important to ensure that they perform in an identical way to standard false belief tasks. This task was therefore first benchmarked against a classic false belief task (the so-called "Smartie" task) in both normal children aged 3-6 years old and in autistic adults. Autistic humans are a strong test case because one of their defining clinical features is that they lack theory of mind and are unable to solve false belief tasks (LESLIE 1987; BARON-COHEN 1990; HAPPÉ 1994). In order to validate the analogue false belief task, we need to confirm that, while autistics cannot pass the task, children show the classic transition between failing the task at age 3 years and passing it by age 5-6 years. Having done that, we then ask how chimpanzees perform on it.

Methods

Subjects

Four chimpanzees acted as subjects for this study. For the purposes of this experiment, we were allowed to test two female and one male (all aged 4–6 years) at Twycross Zoo (East Midlands, England) and a 22-year-old adult male, Pepe, at Dudley Zoo (West



Figure 1. Schematic diagram of apparatus as seen by the subjects. The apparatus consisted of a wooden box (91.5cm wide, 63 cm high and 46 cm deep) into the lower half of which were set four drawers. The drawers could be baited by the experimenter from the rear of the apparatus. A belt onto which a peg could be hung above one of the drawers was set into the upper part; the experimenter could move the belt (such that the peg moved with it) by turning a lever at the back of the apparatus.

Midlands, England); all had been born in captivity. The chimpanzees at both zoos were housed in groups (of 2 to 7 individuals), with each group having its own complex of individual night cages, indoor area and outdoor enclosure. At the time of the study, one of the Twycross chimpanzees was being housed on her own because of previous attacks by other members of her social group. The indoor and outdoor areas had a variety of enrichment devices (ropes, climbing frames, etc). The Twycross animals were fed in their individual cages morning and evening, with seeds and other dried foods scattered around the indoor cage during the day; at Dudley, the animals were fed together as a group. None of the chimpanzees had any previous experience of experiments that involved the need to manipulate objects, though all the Twycross individuals had been involved in previous observational experiments (e.g., joint attention tasks: see O'CONNELL 1995).

A total of 41 children aged 3-6 years old from a local primary school and 16 autistic adults from a residential home (mental ages 2y6m to 13y10m) were used to benchmark the analogue false belief task. Autistic subjects were first given the British Picture Vocabulary Scale (BPVS) and only those whose language competence was sufficient to understand the verbal instructions for the standard false belief task (the "Smartie" task) were included in the study. Five of the autistic adults were subsequently dropped from the study because they failed to complete the training phase of the experiment successfully.

Evolution and Cognition | 132 | 2003, Vol. 9, No. 2

All human subjects participated voluntarily, and were free to withdraw at any time; they were identified during the experiments only by their first names and these were subsequently anonymised. All experiments with the autistic subjects were carried out under the supervision of Francesca HAPPÉ under the auspices of her own study of these individuals.

Apparatus

The apparatus was a wooden box (91.5 x 63 x 46 cm) with 4 small wooden drawers set into its front face (Figure 1). The drawers could be baited with reward items from the back of the apparatus. A fan belt, upon which a large red wooden "dolly" peg could be hooked, ran along the length of the top of the front face above the drawers. The fan belt could be moved by a lever by an experimenter standing at the back of the apparatus. The face of the apparatus was sufficiently high that subjects seated in front of the apparatus could not see behind its façade or see what the experimenter seated behind the apparatus was doing: they were thus unable to see either which draw was being baited (or who was causing the peg attached to the fan belt to move from a position above one drawer to a position above another when this was done remotely by the experimenter).

Procedure

All subjects were tested individually on their own out of sight of the other subjects. The apparatus was placed in front of the subject (but out of reach). The experimenter placed a peg above one of the drawers, hooking it onto the fan belt so that it stood upright. Having then shown the subject the reward (a piece of fruit or a sweet), she then went round the back of the apparatus and placed the reward in the corresponding drawer. From the front of the apparatus, the subject could see the experimenter above the top of the apparatus but not what she was doing. Once the target drawer had been baited, the apparatus was moved up to the subject so that it could choose a drawer to open (by pulling on the appropriate drawer knob). If the subject had chosen the baited drawer, it was allowed to remove the reward. After a short time-out period, the procedure was repeated, with the next drawer to be baited being chosen at random. In all except the familiarisation trials (see below), the subject was allowed only one choice of drawer, the apparatus being moved out of reach immediately after the choice had been made. In all cases, an assistant remained with the subject at the front of the apparatus to control his/her access to the drawers so as to prevent both premature and multiple choices. For the chimpanzees, SO'C acted as the experimenter and a keeper as the assistant; SO'C acted as the assistant for all the human studies, with a teacher as the experimenter for the children and Francesca HAPPÉ in this role for the autistic subjects.

Initially, subjects were given a series of familiarisation trials in which they were allowed to open as many drawers as they liked until they found the reward. However, once they had understood both how to open the drawers and that there would be a reward in only one of the drawers, they were allowed to make only one choice on each trial. Training was terminated when the subject achieved an overall response rate that was significantly above chance. This was taken to be four correct choices in succession for the children or six correct choices out of a block of 10 trials for the autistic subjects and the chimpanzees. A more lenient criterion was allowed for the autistics and chimpanzees because of their slower and less certain rate of learning. The children required only 4-8 training trials to reach criterion, whereas the autistic adults required 20 and the chimpanzees 42-117 trials.

Once the training phase was complete, each subject was presented with the test phase. Test trials were identical to training trials, except that, after the experimenter had placed the indicator peg on the belt and moved behind the apparatus, the peg moved (or was moved by the assistant) to a position above another drawer. The procedure differed slightly for the different groups of subjects: for the children, the autistics and one chimpanzee (Pepe at Dudley Zoo), an assistant moved the peg by hand in full view of the subject (but out of the experimenter's view), but for the remaining chimpanzees the experimenter moved the peg from behind the apparatus by means of a lever.

The task relies on the subject having learned that it is the experimenter who baits the drawers and that she reliably baits the drawer marked by the peg; the training phase was designed to ensure that the subject recognised this to be the case. On the test trials, the subject has to decide whether to choose the drawer above which the peg stands at the point when allowed to make a choice or the drawer above which the experimenter placed the peg (indicating that drawer she would bait). The subject thus has to appreciate that the correct response on the training trials is no longer the correct response on the test trials. To choose the correct drawer, the subject has, in effect, to believe that the experimenter has a false belief about which drawer is currently identified by the peg.

Evolution and Cognition | 133 | 2003, Vol. 9, No. 2

Two types of trials (true belief and false belief) were presented in random order. In true belief trials, the experimenter came back from behind the apparatus (reward still in hand) to look at the position of the peg before returning to bait the drawer; in false belief trials, she did not. In true belief trials, the experimenter knew the peg had moved and therefore baited the drawer now marked by the peg; in false belief trials, the experimenter did not know that the peg had moved and therefore baited the drawer that she had originally marked with the peg. Each subject was given the same number of test trials he/she had received in the training phase, with half of the trails being true belief and half being false belief.

True belief trials act as a control for learning: with repeated false belief trials in the test phase, subjects could in principle achieve high performance after the first couple of test trials simply by rapidly learning a new response ("choose the old peg position not the current one"). With a mixture of true and false believe trials, the subject has to be able to differentiate between cases where the experimenter knows the peg has moved (true belief condition) and those where she does not (false belief condition) and must be able to switch correctly between two different response rules from one trial to the next. Subjects that lack theory of mind should be able to perform correctly on true belief trials even though they fail the false belief trials. Subjects that have theory of mind should perform equally well on both types of trial.

In addition to the mechanical false belief task, both sets of human subjects were given a version of PERNER/LEEKAM/WIMMER (1987) "Smartie" false belief task. This task was used to benchmark performance on the mechanical analogue task. In the "Smartie" task, subjects were shown a tube of Smarties (similar to M&Ms) and asked what they thought was inside. When they answered "Smarties", the cap was removed to reveal dried beans. The cap was replaced and the subject was then asked what he/she had thought was in the tube before it was opened. Those who answered "beans" were deemed to have failed the task, those who answered "sweets" or "Smarties" were deemed to have passed (i.e., understood that they had previously held a false belief about the contents). On standard false belief tasks like the "Smartie" task, children normally show a rising improvement in performance with age, with 3-year-olds performing at about chance level and 5and 6-year-olds performing at around 100% correct responses.

Finally, a second test was conducted on the autistic subjects 6 weeks after they had completed the false belief tasks. The aim of this test was to confirm that the reason the autistic subjects failed the false belief task was not due to an inability to learn a response rule but rather to their inability to understand another individual's mental state. This task was identical to the nonverbal mechanical analogue of the false belief test but lacked the social component. Subjects were taught a conditional rule in which the choice of apparatus depended on a particular cue (a large piece of coloured cardboard which ran the whole length of the apparatus placed on the apparatus between the placing of the peg and the baiting of the drawer). One piece of cardboard was red with a large number 1 on it: this indicated that rule 1 was to be followed (namely, that the reward would be in the drawer that had been originally labelled with the peg). This was the equivalent of the false belief condition, but without the need for theory of mind. The second piece of cardboard was silver with a large number 2 on it: in this case, the rule was that the reward would be in the drawer currently labelled by the peg (equivalent to the true belief task but without the need for social understanding). Both experimenter and assistant remained in front of the apparatus until after the cardboard marker had been placed on it. Before the experiment was run, it was established that each subject could correctly tell the difference between the colours and the numbers.

Results

We first need to establish that the mechanical analogue task is a reasonable approximation to a standard false belief task by showing that both the children and the autistic adults perform equivalently on the analogue and Smartie tasks. For these purposes, 3-year-old children and autistics should perform at chance level on both false belief tasks, but 6-year-olds should have little difficulty with the tasks, while 4- and 5-year-olds should pass on the false belief tasks but not perform as well as the 6year-olds. In contrast, all groups of human subjects should perform well on the true belief version of the analogue task because the true belief version can be solved successfully with first order intentionality (both experimenter and subject happen to share the same belief about the state of the apparatus, so that the task can be solved correctly by the subject acting on its own belief-state).

		"Smartie" test a		True belief ^b		False belief ^b	
		correct (%)	Ν	correct (%)	Ν	correct (%)	Ν
Children:	3-yr-olds	40.0	10	65.0	9°	27.5	9°
	4-yr-olds	63.6	11	77.3	11	38.6	11
	5-yr-olds	60.0	10	60.0	10	22.5	10
	6-yr-olds	100.0	10	72.5	10	80.0	10
Autistics		20.0	5 ^d	56.8	11	14.8	11
Chimpanz	ees	N/A		20.0	4	42.5	4

Table 1. Mean percent correct scores on the three tasks for each category of subject. (a) one trial each per subject; expected number of correct scores if choosing randomly is 50%. (b) 4 trials of each per subject for the children, 8 for the autistics and 10 for the chimpanzees, with mean percent correct score for all subjects; expected with random choice is 25%. (c) one child failed to respond and was excluded; one child was given 9 trials. (d) 6 subjects failed to respond and were excluded.

Benchmarking the Analogue False Belief Task

Table 1 summarises the main results for the various tests and these are presented graphically in Figure 2. Note that chance response levels are 50% correct responses on the "Smartie" task (a two-choice task) and 25% correct responses on the analogue task (a four-choice task). The performance of the children and the autistic adults on the standard false belief task (the "Smartie" task) is broadly in line with conventional findings: three-year-old children and autistic adults perform at chance level on false belief tasks, but, as they increase in age, normal children eventually acquire theory-of-mind, such that by age six years they have little trouble with these tasks. However, in this sample only the 6-year-olds scored



Figure 2. Percentage of correct scores on the conventional false belief task (Smartie test) and mean percentage of correct scores on the true and false belief versions of the mechanical analogue task. The solid horizontal line marks the chance level of performance on the mechanical analogue task with four choices (25% correct). The Smartie task is a conventional two-choice false belief task (expected chance performance of 50% correct responses: dashed line). All subjects received one trial on the Smartie task, but four (children), 8 (autistics) or 10 (chimpanzees) trials on each version of the mechanical analogue task.

at significantly better than chance levels on the "Smartie" task (binomial test with parameter p = 0.5: P = 0.001 1-tailed). The 5-year-olds typically performed at a much poorer level on this task than would normally be expected.

Comparison of these results with those for the mechanical analogue false belief task confirms that the performance of the children and the autistic adults on this task parallels that for the standard false belief task (including the unusually poor performance of the 5-year-olds). The only important difference here is that the children clearly found the mechanical analogue a good deal harder to solve than the conventional task: their performance remains at chance level for much longer before showing the expected improvement in performance. This suggests that the mechanical analogue was more demanding: it had four choices, whereas conventional false belief tasks have only two. With this caveat, the results suggest that the mechanical false belief task is a reasonable analogue of the standard false belief task.

Note that both the children of all ages and the autistic adults are broadly competent on the true belief version of the mechanical analogue task, as they should be when they can see that the experimenter knows about the current state of the apparatus. Once again, however, the four-choice structure to the task clearly taxes even the abilities of the sixyear-olds: they fail to choose correctly on about a quarter of trials, even though they can pass more conventional two-choice false belief tasks without error.

One concern might be that the autistic subjects' inability to solve false belief tasks correctly could be attributed to their relatively low verbal ages. However, only four had verbal ages below 5 years (by which time normal children are competent at false

Evolution and Cognition | 135 | 2003, Vol. 9, No. 2



Figure 3. Mean (SD plus 95% range) of percentage correct responses by autistic subjects on the non-mentalised conditional rule experiment. Rule 1 is the equivalent of the false belief task; Rule 2 the equivalent of the true belief task. Sample size is 11 subjects. Each subject received 8 trials on each task except for two who received 16 and one who received 22 and 11 trials, respectively.

belief tasks). Thus, it seems unlikely that their poor performance on both the Smartie task and the mechanical analogue task could be attributed entirely to such a confound.

Finally, the autistic subjects were given a two-rule conditional task in order to determine whether they performed better on a non-mentalising version of the mechanical analogue task. The results are shown in Figure 3. None of the 11 subjects performed at better than chance levels on Rule 1 (the equivalent of the false belief task), but seven did so on Rule 2 (the equivalent of the true belief task). Taken as a whole, the subjects' performance was significantly better than chance on Rule 2 (8/9 with more than 25% scores correct, with two ties: P = 0.02 1-tailed) and they clearly did significantly better on Rule 2 than on Rule 1 ($2 \times 2 \chi^2 = 53.35$, df = 1, $P \ll 0.001$). These results suggest that the autistic subjects were sticking to the rule that the reward was in the draw currently marked by the peg irrespective of whether the peg had moved and were unable to learn a conditional rule. In other words, they were failing the false belief task either because they were not able to understand that the experimenter had a mental state different to their own or because they could not learn a conditional rule to solve the problem that would bypass the need to engage in mentalising strategies, or both.

	t*	df	Р
(a) True belief trials:			
3-year-olds	5.29	7	0.001
4-year-olds	7.35	10	< 0.001
5-year-olds	2.59	9	0.029
6-year-olds	6.04	9	< 0.001
Autistics	3.68	10	0.004
Chimpanzees	-1.26	3	0.308
(b) False belief trials:			
3-year-olds	0.55	7	0.598
4-year-olds	1.15	10	0.277
5-year-olds	-0.23	9	0.823
6-year-olds	5.66	9	< 0.001
Autistics	-1.94	10	0.082
Chimpanzees	7.00	3	0.006

Table 2. Statistical significance of correct response scores on the mechanical analogue task for the different groups of subjects under the two conditions.

*Comparison against an expected value of 25% correct assuming that subjects chose a drawer at random.

Chimpanzee performance on false belief tasks

Having established that the analogue task is a reasonable model for a false belief task in humans, we can now ask how well the chimpanzees perform. When only the first 10 trials of each subject are considered, the chimpanzees' performance on the false belief versions of the mechanical analogue was not outstanding. However, examination of the data in Figure 2 suggests that their overall performance on the false belief task was at least as good as that of the 4-5 year old children. Table 2 compares scores on the true belief and false belief versions of the mechanical analogue task against the expectation that subjects chose drawers at random (expected success rate of 25%). These results show that all the humans (including the autistics) score at significantly better than chance on the true belief version of the task, but chimpanzees do not; on the false belief version, only the six-year-olds and the chimpanzees score at rates significantly above chance levels.

An analysis of variance (with percentage correct on test trials as the dependent variable) indicates that there are significant differences between the categories of subjects on both types of task (true belief: $F_{5,48} = 2.83$, P = 0.026; false belief: $F_{5,48} = 6.52$, P < 0.001). (Both true and false belief response rates are normally distributed: one-sample KOLMOGOROV-

Evolution and Cognition | 136 | 2003, Vol. 9, No. 2


Figure 4. Percentage of choices made for the previous and current peg positions and the other two positions (combined) in the mechanical analogue task. Top: True belief version. Bottom: false belief version.

SMIRNOV tests, P = 0.104 and P = 0.083 respectively; however, as a check, we ran a non-parametric ANOVA, but the results were the same: P = 0.042and P = 0.003 respectively.) Bonferroni *post hoc* tests confirm that six-year-olds score significantly higher ($P \le 0.027$ 2-tailed) than all other groups of subjects except the chimpanzees (for whom P = 0.48 2tailed) on the false belief tasks, but not on the true belief tasks (except against the chimpanzees, where P = 0.058 2-tailed). Conversely, the chimpanzees differ significantly ($P \le 0.058$ 2-tailed) from all the human subjects (except the five-year-olds and autistic subjects, for whom $P \ge 0.38$) on truth belief tasks, but not on the false belief tasks.

An alternative way of considering the results is to compare the number of occasions when subjects selected the previous and current positions of the pegs in the true belief and false belief trials. Subjects should choose the current position in true belief trials and the previous position in false belief trials. Figure 4 plots the mean percentage of choices for the four drawers (summing the scores for the two "other" choices). It is clear from these results that only the 6-year-olds showed the correct switch from current to previous peg position between the two types of trials. Younger children and the autistic adults all exhibited a strong tendency to select the previous position of the peg in both types of trial. Significantly, perhaps, chimpanzees exhibited a weak tendency to switch choices correctly with trial type, even though they made many more incorrect decisions ("other" drawers) than any of the human subjects. However, only one chimpanzee (Josie) selected the previous peg position on false belief trials and the current one on true belief trials significantly more often than expected by chance ("other" choices discounted: 2×2 $\chi^2 = 21.88$, df = 1, P < 0.001). Matched pairs *t*-tests confirmed that, as a group, only the 6-year-olds chose the previous peg position significantly more often in the false belief trials (t = 2.42, df = 9, p < 0.05) and the current peg position significantly more often in the true belief trials (t = 2.87, df = 9, p < 0.02); all other comparisons were non-significant. This suggests that, even though their performance was far from perfect, the chimpanzees had a better grasp of the false belief task than either the autistic adults or the younger children.

Learning and task performance in chimpanzees

The chimpanzees were given extended trials on both the true belief and false belief tests in order to ascertain the extent to which learning might play a role in their responses to the tasks. Regression analyses of the results using successive blocks of 10 trials on each test type are summarised in Table 3. Pooling results using Fisher's procedure (with 1tailed *P*-values in a positive direction to test for an underlying trend for scores to improve with time) yields $\chi^2 = 4.56$ (df = 2 × 4 = 8, *P* = 0.80) for the true belief task and χ^2 = 7.35 (df = 8, *P* = 0.50) for the false belief task, indicating that there is no underlying trend in either case. Indeed, the mean slope on

Subject	slope	r2	t	df*	Р
(a) True belief trials:					
Рере	-0.004	0.00	-0.035	11	0.959
Josie	0.600	0.14	0.707	4	0.530
Flynn	-0.100	0.01	-0.141	4	0.897
Beckie	0.114	0.01	0.231	4	0.829
(b) False belief trials:					
Рере	-0.025	0.01	-0.273	11	0.717
Josie	0.600	0.53	0.187	4	0.164
Flynn	-0.800	0.46	-1.589	4	0.210
Beckie	-0.400	0.38	-1.572	4	0.191

Table 3. Regression statistics for a linear trend in correct scoreswith blocks of trials for chimpanzees to test for a learning effect.*blocks of 10 test trials.

Evolution and Cognition | 137 | 2003, Vol. 9, No. 2

the false belief tasks is negative (b = -0.156), while that for true belief tasks is slightly positive. This suggests that learning is not giving rise to a disproportionate level of positive results with the longer trial sequences in the chimpanzees.

Discussion

Although these results do not unequivocally suggest that chimpanzees have theory of mind in the strong sense that we would normally make that claim of 5- and 6-year-old children, they do indicate that chimpanzees perform at higher levels on false belief tasks than autistic adults and 3-year-old children, both of whom are known not to possess theory of mind. One conclusion we might draw is that chimpanzees are hovering on the same threshold as children (aged around 4 years) who are in the process of acquiring formal theory of mind. That having been said, however, it is quite apparent that the chimpanzees operate at a much lower level than even those children with whom they might be considered to be cognitively comparable: they took far longer to learn the task to criterion during training (42–117 trials) than either the children (up to eight trials) or the autistic adults (20 trials).

Nonetheless, we should, perhaps, be cautious about interpreting these results as evidence for theory of mind (however weak) in chimpanzees. The fact that, unlike all the human subjects, the chimpanzees (just about) passed the false belief tasks but failed the supposedly less taxing true belief tasks raises doubts as to how they were making their decisions. A plausible interpretation of these results might be that the chimpanzees had learned a new rule on test trials ("choose the original peg position, no matter what"), and were unable to factor into this the observation that the experimenter had seen the change in peg position in the true belief trials. This interpretation would imply that chimpanzees did not appreciate the relationship between seeing and knowing on the part of the experimenter, and the results should therefore be interpreted as providing evidence that they categorically do not possess theory of mind. However, this conclusion would conflict with HARE et al.'s (2000; HARE/CALL/TOMASELLO 2001) experimental results showing that chimpanzees can successfully pass "seer-knower" tests. It is also at odds with the finding that the chimpanzees (but not the younger children or the autistics) did show some evidence of switching appropriately between true and false belief tasks (Figure 4).

An alternative concern is the possibility that the experimental design might allow the analogue task to be solved without any need for theory of mind. Although every effort was made during the training trials to ensure that all subjects understood that it was the experimenter who baited the apparatus, it is impossible to guarantee that individual subjects were aware of this. It is therefore always possible, in principle at least, that the chimpanzees tried to solve the task using a simpler rule (such as "ignore the fact that the peg has moved in test trials") or that they were able to learn during the first few test trials that they had to use a conditional rule that switched between two responses ("choose the drawer currently identified by the peg" to "choose the drawer that was previously identified by the peg") on the basis of a single circumstantial cue (whether or not the experimenter checked the position of the peg before baiting).

The first possibility can be ruled out as a general explanation because at least one of the chimpanzees (Josie) exhibited a significant tendency to switch responses appropriately between true and false belief trials (Figure 4). The second possibility is more difficult to deal with. It is, however, significant that the chimpanzees and autistic subjects responded in very different ways to the switching between true and false belief tasks. Figure 4 suggests that the autistic subjects (in common with the 3-year-olds) were unable to break away from the simple rule "choose the drawer that is currently indicated" (hence, their good performance on true belief tasks and poor performance on false belief tasks). In contrast, the chimpanzees, though far from perfect, did show a tendency (albeit significant only in the case of Josie) to switch correctly between trial types, a pattern that is also apparent among the 4- and 5year-old children. Note that, among the children, only 6-year-olds demonstrated unequivocal competence in this respect. More importantly, perhaps, if this explanation was to hold, we would have to assume that the chimpanzees are, in effect, capable of switching their learning competences at will: during the training trials, they required 80–100 trials to learn the simple rule that the peg indicates which box is baited, yet a few days later were all suddenly capable of learning a complex conditional rule on the basis of at most 10 test trials. Such a dramatic switch in competence is implausible and cannot really be taken seriously as an alternative explanation to the simpler interpretation that the chimpanzees show evidence of some (albeit low grade) mindreading.

There is, in addition, another reason for be cautious about concluding, on the basis of these results alone, that chimpanzees do not possess theory of mind. Our analogue task was clearly more complex than the conventional two-choice tasks that are normally used in studies with children: it certainly proved

taxing to the younger children and extremely taxing to the autistic adults. Although chimpanzees also failed the two-choice false belief task used by CALL/TOMASELLO (1999), it may be that the kinds of false belief tasks used in these study are simply too difficult or too anthropocentric for chimpanzees to understand, and are thus unfair tests of the chimpanzees' mentalising capabilities. This interpretation is given some support by HARE et al. (2000; HARE/CALL/TOMASELLO 2001), who showed that chimpanzees can pass more naturalistic tasks that involve understanding the visual perspective and likely knowledge state of another individual.

One residual concern might be the different numbers of test trials that the various categories of subjects received. Autistic adults received twice as many trials as the normal children, while the chimpanzees received more than ten times as many. The principal concern here must be that subjects might learn an appropriate rule for the correct response if given enough trials. Two considerations militate against any such a concern, however. First, subjects were given true and false belief trials in randomised order, and thus had to be able to switch, often in successive trials, between one response and the other. To have been able to respond at better than chance despite this interference itself suggests a significant level of cognitive competence. Second, the autistic subjects failed the false belief tasks despite repeated exposure. Indeed, while they were able to learn a simple rule ("choose the apparatus currently marked by the peg"), they were unable to learn a two-rule task, suggesting that the problem lay not with an inability to learn *per se* but with an inability to learn the complex conditional rule needed to solve the unpredictable switch between true and false belief trials.

The chimpanzees' slowness in learning the task during the training phase compared even to 3-yearold children remains enigmatic, though perhaps far from unusual in studies of this species. Whether this

Authors' address

Sanjida O'Connell, BBC Natural History Unit, Broadcasting House, Whiteladies Road, Bristol BS8 2LR, UK, and R.I.M. Dunbar, Evolutionary Psychology Research Group, School of Biological Sciences, Crown St, University of Liverpool, Liverpool L69 3GS, UK. Email: rimd@liv.ac.uk implies that chimpanzees are significantly poorer learners than human children or are simply less willing to apply themselves remains unclear. The tasks were presented as a game to both the children and the autistic adults, and children are invariably enthusiastic about any form of game. Devising tasks that are equally

interesting for both species remains a challenge that has not yet been overcome by any study. Indeed, in most studies involving autistics (this study included), it is necessary to structure the task in a relatively lenient way in order to engage the subjects at all. Six autistic subjects declined to respond at all in the Smartie test, for example; as a result, the autistics who went on to do the mechanical analogue task were given exaggerated cues by the experimenters to reinforce the looking and deception components. Thus, if anything, the autistic subjects may have benefitted from cuing advantages that were certainly denied to the chimpanzees. The fact that the chimpanzees did significantly better than the autistic subjects despite such a bias merely reinforces their better performance.

On balance, then, we conclude that chimpanzees have demonstrated some, albeit weak, competence on theory of mind tasks. By this, we mean that they are able to perform at about the same level as children who are on the verge of acquiring formal theory of mind. However, because the chimpanzees performed unexpectedly poorly on the true belief tasks—and we have no really adequate explanation as to why they might have done so—we remain hesitant to claim that they have demonstrated full competence in this respect.

Acknowledgments

We thank the head and staff of St Anne's RC First School, Nuneaton, the manager and staff of SANDs, Gravestone, Kent, and Molly BADHAM and her staff at Twycross Zoo for permission to carry out the study at their respective establishments. Nadine WILLIAMSON and Erica LONG kindly acted as experimental assistants. We are particularly grateful to Dr Francesca HAPPÉ for giving us access to the autistic subjects from her study and for helping run the experiments with the autistic subjects. SO'C was supported by a UK BBSRC postgraduate studentship.

References

- Astington, J. W. (1993) The Child's Discovery of the Mind. Cambridge University Press: Cambridge MA.
- Baron-Cohen, S. (1990) Autism: A specific cognitive disorder of "mind-blindness". International Review of Psychology 2: 79–88.
- Byrne, R./Whiten, A. (1991) Computation and mindreading in primate tactical deception. In: Whiten, A. (ed) Natural Theories of Mind. Blackwell: Oxford, pp. 127–142
- Call, J./Tomasello, M. (1998) Distinguishing intentional from accidental actions in orang utans (Pongo pygmaeus), chimpanzees (Pan troglodytes) and human children (Homo sapiens). Journal of Comparative Psychology 112: 192–206.
- Call, J./Tomasello, M. (1999) A nonverbal theory of mind test: the performance of children and apes. Child Development 70: 381–395.
- Cheney, D. L./Seyfarth, R. M./Palombit, R. A. (1996) The function and mechanisms underlying baboon contact barks. Animal Behaviour 52: 507–518.
- Happé, F. (1994) Autism: An Introduction to a Psychological Theory. University College London Press: London.
- Hare, B./Call, J./Agnetta, B./Tomasello, M. (2000) Chimpanzees know what conspecifics do and do not see. Animal Behaviour 59: 771–785.
- Hare, B./Call, J./Tomasello, M. (2001) Do chimpanzees know what conspecifics know? Animal Behaviour 61: 139–151.
- Hare, B./Call, J./Tomasello, M. /Visalberghi, E. (2003) Do capuchin monkeys, Cebus capella, know what conspecifics do and do not see? Animal Behaviour 65: 131–142.
- Hauser, M./Wrangham, R. W. (1987) Manipulation of food calls in captive chimpanzees. Folia Primatologica 48: 207– 210.
- Heyes, C. M. (1993) Anecdotes, training, trapping and triangulating: do animals attribute mental states? Animal Behaviour 46: 177–188.
- Heyes, C. M. (1998) Theory of mind in nonhuman primates. Behavioral and Brain Sciences 21: 101–148.
- Leslie, A. M. (1987) Pretence and representation in infancy: the origins of "theory of mind". Psychological Review 94:

84-106.

- **O'Connell, S. (1995)** Theory of mind in chimpanzees. PhD thesis, University of Liverpool.
- Perner, J./Leekam, S. R./Wimmer, H. (1987) Three-year-olds' difficulty with false belief: the case for a conceptual deficit. British Journal of Developmental Psychology 5: 125–137.
- **Povinelli, D. J./Eddy, T. J. (1996)** Chimpanzees: Joint visual attention. Psychological Science 7:129–135.
- Povinelli, D. J./Nelson, K./Boysen, S. (1990) Inferences about guessing and knowing by chimpanzees. Journal of Comparative Psychology 104: 203–210.
- Povinelli, D. J./Parks, K./Novak, M. A. (1991) Do rhesus monkeys (Macaca mulatta) attribute knowledge and ignorance to others? Journal of Comparative Psychology 105: 318– 325.
- Povinelli, D. J./Parks, K./Novak, M.A. (1992) Role reversal in rhesus monkeys, but no evidence for empathy. Animal Behaviour 44: 269–281.
- Povinelli, D. J./Perilloux, H./Reaux, J./Bierschwale, D. (1998) Young chimpanzess' reactions to intentional versus accidental and inadvertent actions. Behavioural Processes 42: 205–218.
- **Premack, D./Woodruff, G. (1978)** Does the chimpanzee have a theory of mind? Behavioral and Brain Sciences 4: 515–526.
- Rendall, D./Cheney, D. L./Seyfarth, R.M. (2000) Proximate factors mediating 'contact' calls in adult female baboons (Papio cynocephalus ursinus) and their infants. Journal of Comparative Psychology 114: 36–46.
- Seyfarth, R. M./Cheney, D.L. (1990) The assessment by vervet monkeys of their own and another species' alarm calls. Animal Behaviour 40: 754–764.
- Tomasello, M./Call, J./Hare, B. (1998) Five primate species follow visual gaze of conspecifics. Animal Behaviour 55: 1063–1069.
- Whiten, A./Byrne, R. (1988 Tactical deception in primates. Behavioral and Brain Sciences 11: 233–273.
- Wimmer, H./Perner, J. (1983) Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. Cognition 13: 103– 128.

Was the First Craniate¹ on the Road to Cognition?

A Modern Craniate's Perspective

ne might propose three mutually exclusive hypotheses to explain the existence of human-level cognition² observed on Earth:

1. Human-level cognition occurs as a *fluke* in the universe; on Earth it may be explained as an incidental result of highly contingent events.³

2. Human-level cognition is a common occurrence throughout the universe; its existence on Earth may be explained as a result of highly constrained or law-like processes.

3. Human-level cognition on Earth can be explained as neither fluke nor the result of law-like processes;⁴ the frequency of its occurrence in the universe may not be predictable.

An investigator who

knew that hypothesis 1 or 2 was correct would be able to comment knowledgeably upon the probability of human-level cognition existing on any planet where signs of life were observed. To estimate the probability of human-level cognition on such a planet, a scientist might only need to know the length of time that the planet had sustained conditions for life. Ideally, an investigator attempting to choose between the above hypotheses should ob-

serve life at several stages on many habitable planets

Abstract

Chinese fossil discoveries of the earliest known craniates (from the early Cambrian period) have led scientists to question whether the evolution of human-level cognition is a rare occurrence in the universe. The earliest chordate is now best represented by a well-documented metazoan called Haikouella lanceolata. Possessing a relatively large brain, this animal appears to demonstrate that the brain and endoskeleton did not evolve together, as had been assumed, but rather that the brain appeared long before full endoskeletization. The paleontologist who describes the animal further notes a "top-down" pattern in the appearances of new forms in the fossil record. Researchers find such observations relevant to the question: Was the evolution of human-level cognition in some sense inevitable, or was it an accident dependent upon historical contingencies? The new evidence for early craniates lends support to the view that human-level cognition may be part of a developmental package, but historical contingencies pose serious problems for a strictly law-like explanation.

Key words

Contingency, convergence, developmentalism, directionality, internalism, saltation, top-down evolution. before claiming confidence in his selection. Of course, the best an Earthbound scientist can do at present is to search for helpful clues on the one planet where we know that life exists; and here at least Earth's crust does provide helpful access to pertinent glimpses of past life stages. Of greatest relevance for any such researcher is the fossil record's snapshot in which the body plan⁵ first appears that would eventually house human-level cognition.

To join such an investigation, we would like to explore the following questions: When does this body plan appear relative to other body plans? Does this body plan have an assured, or at least a likely, survival to permit time for the evolution of

cognition—or is the survival of its future lineage dependent upon highly contingent events? Do other body plans appear that would also seem to make good candidates for the support of human-level cognition? How predictable is the appearance of a suitable, large brain-carrying body plan? For that matter, upon what basis can we predict body plans at all, or changes in any form of life? Does evolution proceed in a bottom-up fashion where small changes accrue into large ones aided by no higher

Evolution and Cognition | 141 | 2003, Vol. 9, No. 2

law than natural selection, or are there top-down forces at work, driving organisms more forcefully toward obligatory forms?

Examining these questions takes us back to the early Cambrian period, to the time, 530 million years ago, when virtually all the major animal groups, called phyla, first appear in the fossil record. In December 1999, the journal Nature announced the discovery there of what appears to be the remains of our earliest chordate ancestor (CHEN/HUANG/LI 1999a). Found in southern China's Yunnan Province near Kunming, the stiffspined, paper clip-size animal named Haikouella lanceolata pushes back the evolution of our own phylum, Chordata, to the very start of modern metazoan life. Until its discovery, paleontologists had begun to despair of ever finding undisputed evidence that our own "advanced" phylum existed as early as the main burst of the Cambrian "explosion" of new body plans. The plentiful fossil evidence for Haikouella-305 specimens, many in excellent condition-finally confirms what many had suspected in recent years: Our own phylum arrived on the scene along with most of the others, during the surprisingly quick radiation of new major animal groups that characterizes the early Cambrian period.

The discovery demonstrates that our chordate ancestors had to fight their way through the rough and tumble Cambrian seas, in a period when many animal groups that entered did not survive to the end. Most families-and even many whole body plans-disappeared (GOULD 1989, pp47n; WARD 2000, p184). As we detail below, many paleontologists further classify Haikouella as a craniate. Any argument for the inevitability of craniate survival could be made more confidently if this body plan had arisen after the Cambrian, since the Cambrian represents the only period in which the number of animal phyla actually decreased (Dobzhansky et al. 1977, pp. 422–423; SOLÉ/GOODWIN 2000, pp. 249– 251). Thus, as we will see, the appearance of the craniate body plan in the early Cambrian has paradoxical implications for a heady future; it can be taken either as evidence for a head start on the pathway to intelligence, or evidence of heightened probability that this lineage would be cut off-at its head. Haikouella simultaneously creates support for two competing inferences: the evolution of cognition as a highly constrained, or as a highly contingent, process.

Current Controversies

The animal's discovery thus raises questions at the heart of current controversies in evolution research. One of the broadest ways to characterize the competing positions is as a disagreement between externalists and internalists, i.e., those who treat external selection as virtually the sole creative force in evolution versus those who emphasize the importance of internal constraints. Related issues tend to pit developmentalists against neo-DARWINISTS, formalists against functionalists, punctuationalists against gradualists, and top-down theorists against bottom-up theorists, with the first party in each pair siding with the internalists.⁶

Perhaps the most profound movement in the field in recent decades has been the advancement of developmental biology (which studies the way genes control the growth of individual living organisms) as a key to evolutionary biology. Since the mid-1990s, the marriage between the two disciplines has become known as evolutionary developmental biology, or simply "evo-devo". Evo-devo explores how changes in ontogeny (the development of individual organisms, from fertilized egg to maturity) are related to the emergence of new phenotypes over successive generations. The goal is to use knowledge of how genes control the development of embryonic structures to learn how these same genes were involved in the first appearance of such structures in past epochs. Early contributors to the field include DE BEER (1930), SCHMALHAUSEN (1949), WADDINGTON (1957), RAFF/KAUFFMAN (1983), ARTHUR (1988), and HALL (1992). Developmentalists believe that too much credit has been given to the power of natural selection.

Wallace ARTHUR finds it "strange ... that mainstream neo-DARWINIAN theory has come to regard natural selection as the primary mechanism causing evolutionary change. Selection is a destructive force, which acts only to eliminate" (ARTHUR 1997, p241). Though trained as a neo-DARWINist, ARTHUR acknowledges that selection "does not create the new type in the first place", adding: "We have come to accept a theory of evolution that explains the origin and diversification of exquisitely engineered organisms on the basis of the selective destruction of genetic/developmental variants whose initial production has been treated, for the most part, as a 'black box'" (Ibid). When it comes to the origin of body plans, according to ARTHUR, explaining it "purely in selective terms, without reference to the underlying genetic architecture will ultimately fail.

Hence the need for the new discipline of evolutionary developmental biology" (p. 291).

Harvard zoologist Stephen Jay GOULD, who died in 2002, noted that DARWIN himself had not believed that natural selection was the exclusive means of evolutionary modification. Nevertheless, he said, certain "ultra-DARWINISTS" were trying to "out-DAR-WIN DARWIN" by claiming "that natural selection regulates everything of any importance in evolution" (GOULD 1997). For their part, neo-DARWINISTS accused GOULD of wandering too far from the reductionism⁷ so necessary to science's success. Daniel DENNETT even suggested that GOULD must have had a "hidden agenda" to sneak purpose back into biology (DENNETT 1996). The accusation was patently false; GOULD never wrote with any such intent; in fact, he went to great lengths to create new terms to replace any that smacked of teleology. "Preadaptations" became "exaptations" precisely to avoid the teleological overtones; "saltation" and "laws of form" became "facet-flipping" (GOULD 2002, pp342-351), helping to assure that the element of chance overshadowed any Platonic connotations.

In his last monograph, GOULD wrote: "I argue that 'internalism' poses two separate challenges to pure DARWINIAN functionalism: saltational change arising from internal forces of mutability, and inherent directionality of variation... Most internalists ... emphasize the second theme of channels and preferred directionality of variation" (2002, p445). One of the traits that made GOULD stand out from other internalists was his emphasis on chance. He seemed intent to exonerate himself from any charge of advocating teleology by assuring us that, though constraints have their important place, stochastic events have the final word. Our own existence is attributable to a "golden happenstance" in the Cambrian explosion (p. 1159), the start of an unlikely course that continued with the repeated overcoming of odds to produce humans as "an ultimate in oddball rarity" (1996-1997). If the developmental patterns of bilaterians appear to have become fixed into "limited and excellent, perhaps even optimal, designs", it is only because they represent just one possible solution among numerous entirely plausible alternatives of strikingly different form, each yielding a subsequent history of life entirely different from the outcome actually experienced on earth" (2002, p1159). Yet all who support development's importance to evolution (including GOULD, who used terms like "directionality" and "congealed designs") must support the concept of directionality, however they explain it.

ARTHUR uses the term "directionality" to describe the observation of large-scale evolutionary changes preceding smaller modifications, "big 'experiments" giving way to progressively more restricted modifications" (ARTHUR 1997, p207), i.e., the top-down concept that we will explore below. He writes: "This idea of directionality—which is absent from conventional neo-DARWINIAN theory—is important in Evolutionary Developmental Biology" (Ibid).

Against this background, we can now return to the findings about the earliest known craniate, to note the support they yield for this article's first two hypothetical options: human-level cognition as either fluke or common occurrence. We can also compare the support they afford straightforward neo-DARWINISM versus developmentalism.

A listing of *Haikouella*-related supporting evidence for developmental constraints—and the ubiquity of human-level cognition in the universe—could read as follows:

(1) early priority of cephalization over endoskeletization,

(2) constraints/convergence/channeling,⁸

(3) hierarchical phylogenies,

- (4) saltation,
- (5) top-down pattern in the fossil record, and

(6) the principle of mediocrity.

Supporting arguments for contingency—and the rarity of human-level cognition in the universe—would include:

(1) The appearance of craniates in the early Cambrian (the one period where the number of phyla decreases) greatly increases the probability that this lineage would become extinct;

(2) Non-chordate phyla can be shown to be incapable of developing human-level cognition; and

(3) Human-level cognition has evolved only once on this planet.

As we recount the findings reported by *Haik-ouella's* investigators, we will examine each of these in turn.

Settling the Matter of Chordates in the Cambrian

Before Jun-Yuan CHEN's discovery of *Haikouella* (CHEN/HUANG/LI 1999a; CHEN/HUANG/LI 1999b) (Figures 1 and 2), it was beginning to look doubtful that Cambrian chordates would ever become firmly established. Each new interpretation of the mostly poorly preserved would-be-chordate specimens was hotly disputed.

Evolution and Cognition | 143 | 2003, Vol. 9, No. 2

Biologists belonging to the contingency camp, as opposed to those expecting directionality, would have settled for an extremely primitive-looking chordate or half-chordate. Champions of chance expected the earliest chordate to exhibit brawn before brains; they sought the distinguishing chordate feature of muscles arranged in V-shaped blocks (called myotomes) without expecting to find much of a head at first. Even more essential for an earliest chordate candidate is the notochord, from which chordates get their name. The notochord is a stiffening, muscularized rod that runs down the middle of the back. The head was seldom mentioned except in reference to things to come, and never as an essential element for the earliest chordate. To illustrate what the earliest chordate should look like, biologists chose amphioxus, an animal that appears headless and pointy at both ends, as the most plausible living model. The name "amphioxus" literally means "both [ends] pointed".

The first generally recognized evidence for possible chordates in the Cambrian came to light almost a quarter of a century ago, when Cambridge paleontologist Simon CONWAY MORRIS tentatively promoted a middle Cambrian species called Pikaia from annelid worm to chordate status (CONWAY MORRIS/ WHITTINGTON 1979). Stephen Jay GOULD saved the tiny, simplified eel-shaped animal for the climax of his popular book about Canada's Burgess Shale animals. On the last pages of Wonderful Life, he called Pikaia "the missing and final link in our story of contingency-the direct connection between Burgess decimation and eventual human evolution" (GOULD 1989, p322). Since then, however, less enthused scientists have questioned Pikaia's chordate classification because of its lack of chordate features like gills, gonads, and a full notochord (HOLLAND personal communication).

For GOULD, the middle-Cambrian *Pikaia* best fit what the earliest chordate should look like: simple, sleek and headless. He mentions no brain, eyes, or other sensory organs when describing *Pikaia* in his popular book; even the possibility of a head seems remote in an animal whose anterior end, in his illustration, splits into two (GOULD 1989).

The next best hope came from animals whose chordate status was disputed or who appeared too late to show that chordates joined in the early Cambrian radiation of new forms. The eel-like conodont, long known only from its teeth, extended back only to the late Cambrian (PURNELL/DONOGHUE 1997). During the 1990s battles ensued over descriptions of two new chordate claims represented by



Figure 1. One of 305 *Haikouella lanceolata* fossil specimens from Haikou, near Kunming, China (early Cambrian period, 530 mya).



Figure 2. Anatomical interpretation of *Haikouella lanceolata* (gen. et sp. nov.). Abbreviations: Abv (anterior branchial vessel); An (anus); Ap (anterior projection); At (atrio); Atp (atriopore); Ba (branchial arches); Baf (banchial-arch-filamentals); Br (brain); Buc (buccall cavity): Co (copulatory organ); Cp (caudal project); Da (dorsal aorta); Df (dorsal fin); Ds (denticular structure); Eg (endostyle glands); Es (endostyle); Esp (esophagus); Hd (head); Ht (heart); It (intestine); Lb (lobated structures); Le (lateral eye); Mg (midgut); Mm (myomeres); Mo (mouth opening); Ms (myosepta); Mw (median wall); Nc (neural cord); Nt (notochord); Ph (pharyngeal cavity); T (tentacle-like structure); Va (ventral aorta); Vf (ventral fin).

just a few specimens: *Yunnanozoon* (CHEN et al. 1995) and *Cathaymyrus* (SHU/CONWAY MORRIS/ ZHANG 1996). The discoverer of *Cathaymyrus* thought *Yunnanozoon* looked more like a hemichordate (acorn worm) than a chordate (SHU/ZHANG/ CHENG 1996); and the discoverer of *Yunnanozoon* opined that his challenger had mistaken *Cathaymyrus's* squashed dorsal fin for a notochord (CHEN/ HUNAG/LI 1999b).

After the Cambrian waters had been sufficiently muddied, researchers wondered if any true chordate had ever been found in Cambrian strata. Maybe our own "sophisticated" phylum had not yet evolved. Even GOULD's *Pikaia*, though used to illustrate Cambrian chordates in vertebrate textbooks, no longer

Evolution and Cognition | 144 | 2003, Vol. 9, No. 2

looked convincing, since it lacked many of the chordate features claimed by the more recent finds.

Thus the significance of the discovery of *Haik-ouella*—displayed in over 300 specimens. In CHEN's description of *Haikouella* fossils, he pointed out features that not only demonstrate its chordate status, but that shed light on the origin of craniates (biology's new name for vertebrates) (CHEN/HUANG/LI 1999b). The new nomenclature reflects a new primary diagnostic feature for this taxon: a distinct head enclosing a brain and sensory organs, recognizing that this character should now take precedence over the presence of a vertebral column.

Known for his research on amphioxus, the present-day animal thought to best represent the ancestor of all vertebrates, Nicholas HOLLAND said: "There's no question these things are chordates" (ENSERINK 1999). He remarked on the great number of specimens with conspicuous gill slits (for straining food out of the water) and other diagnostic characters: "The muscle segments are unarguable, and the notochord's good too" (HOLLAND 1999). Unlike specimens from other recent finds, both Haikouella and Yunnanozoon exhibit large notochords that clearly run the full length of their bodies. "It's the earliest known chordate ancestor", said HOLLAND. "Every zoology student and every paleontology student for many, many generations is going to have to look at that picture. This is going to be page one, two, three and four of vertebrate texts, and paleontology texts, and invertebrate zoology texts" (HOL-LAND personal communication).

Since the discovery of *Haikouella*, Degan SHU et al. (1999) reported their discovery of two new chordates, Myllokunmingia and Haikouichthys, each based on a single specimen. Collaborator Simon CONWAY MORRIS proposes that the animals had skulls made of cartilage (MONASTERSKY 1999). CHEN notes that the specimens display two important features: distinctive fins (large dorsal and possibly paired ventral fins) and zigzag-shaped segmented muscles, similar to the pattern in modern fish (CHEN, personal communication). Though paleontologists of these various discoveries continue to contend with one another over whose specimens are ancestral to whose---and whose are true chordates---all agree that chordates have now been found in the early Cambrian (ENSERINK 1999; DZIK 1995).

What will happen to GOULD's *Pikaia*, the animal zoology textbooks presently tout as our earliest chordate ancestor? HOLLAND contends that the textbook writers had no business picking up *Pikaia* as a chordate ancestor from GOULD's popular book,

since GOULD was not an authority on the animal (HOLLAND personal communication). GOULD had simply made it fit what he needed to relate the Burgess Shale fauna to humans. "Why do humans exist?" asked GOULD on the last page of *Wonderful Life.* "A major part of the answer, touching those aspects of the issue that science can treat at all, must be: because *Pikaia* survived the Burgess decimation" (GOULD 1989, p323). GOULD had used *Pikaia* to relate *Pikaia* to us and us to his overriding theme: contingency. "What this conference has done", said HOLLAND at the symposium where *Haikouella* was announced, "is to pull the rug out from under *Pikaia*, for sure. Nobody will ever talk about it again" (HOLLAND personal communication).

Shedding Light on Vertebrate Origins

Now that lower Cambrian chordates have been confirmed, zoologists must deal with the fact that Haikouella-and other early Cambrian chordates-look nothing like what they expected to see in a predecessor of Pikaia. Rather than finding evidence that this complex animal had less sophisticated ancestors, CHEN and SHU instead found examples of more complex, fully formed chordates-fifteen million years earlier. None of these newly discovered chordates have vertebrae or endoskeletons, so strictly speaking, they aren't vertebrates. But displaying relatively large brains, these animals appear to be in the line to vertebrates, so that at the conference where Haikouella was announced, the strange term "pre-backbone vertebrate" was frequently bandied about. The brain's early appearance would seem to demonstrate that brain and endoskeleton did not evolve together, as had been assumed, but rather that the brain appeared long before the development of the vertebrate spine.

"The discovery of the first craniate shows that the evolutionary history toward vertebrates had been on track long before the origins of the backbone", says Taiwanese biologist Chia-Wei LI (1999), co-author of the *Haikouella* description. *Haikouella* findings run counter to the commonly-held notion that the head could not become the dominant body structure until the body's superstructure was also in place. It now appears that, against externalist expectations, cephalization (when the head became the dominant or controlling body structure) preceded endoskeletization (the development of an internal support structure).

CHEN also identified other important features in *Haikouella* that preceded the development of a bony

Evolution and Cognition | 145 | 2003, Vol. 9, No. 2

skeleton: a neural cord that, like the notocord, runs the length of the body; a heart; a pair of lateral eyes; and tiny teeth. The teeth are located far back in its large pharyngeal cavity rather than in the mouth, indicating that it used them for grinding, not biting. Biologists had assumed that chordates did not develop the ability to accumulate minerals in their bodies to form teeth or bones until about 500 million years ago. But *Haikouella* and *Yunnanozoon* demonstrate that biomineralization had begun at least 30 million years earlier. Teeth led the way long before the development of a notochord-protecting, mineralized vertebral column or other bones.

Constraints, Channeling, and Convergence

The sudden explosion of widely disparate Cambrian animal Bauplans, followed by no new body plans throughout the rest of geologic history, fits the picture of a constrained process, the channeling of changes within particular forms. Scientists also find evidence of constraints today in the form of parallelism and convergence,⁹ both in experiments with living animals and in theoretical modeling. From his research on the development of amphibians, brain researcher Gerhard SCHLOSSER notes trends "where several characters tend to act as a 'unit of evolution', i.e., they tend to coevolve repeatedly" (SCHLOSSER 2000).

Evolutionary geneticist Paul RAINEY and his colleagues have also noticed convergence in evolution while experimenting with the bacterium Pseudomonas fluorescens. "These experiments in test-tube evolution", says RAINEY, "allow us to replay life's tape, albeit on a small scale, as often as we like" (RAINEY 2003). Their findings? "Evolution repeats itself". By growing rapidly diversifying strains of the bacterium in test tubes of nutrient broth, they have discovered that "in the face of similar selective conditions, different lineages can find similar solutions to the same problems". RAINEY is not afraid to find implications from his findings for human evolution: "Replay life's tape", he claims, "and while Homo sapiens may not evolve there is a high probability that introspective bipedal organisms with binocular vision will" (Ibid).

Simon CONWAY MORRIS reaches a similar conclusion. Speaking of the property of consciousness, he writes: "Here the reality of convergence suggests that the tape of life, to use GOULD's metaphor, can be run as many times as we like and in principle intelligence will surely emerge" (CONWAY MORRIS

1998, p14). What about "the numerous entirely plausible alternatives of strikingly different forms" that GOULD expected if the tape should be rerun from the beginning? "Put simply", says CONWAY MORRIS, "contingency is inevitable, but unremarkable.... There are not an unlimited number of ways of doing something. For all its exuberance, the forms of life are restricted and channeled" (p13). CONWAY MORRIS believes that convergence "effectively undermines the main plank of GOULD's argument on the role of contingent processes in shaping the tree of life" (Ibid). GOULD, he says, "presupposes that constraints are weak" and makes a "most egregious misinterpretation of the Burgess Shale" (CON-WAY MORRIS 1998–1999). His "egregious misinterpretation"-contingency as the major lesson of the Burgess Shale—is a conclusion that GOULD drew from his personal credo, according to CONWAY MOR-RIS, not from paleontology (Ibid).

Hierarchies

Cladistics, a branch of biology that does indisputably draw its evidence from paleontology, hypothesizes relationships between organisms according to shared derived characters (synapomorphies). The distribution of these diagnostic features forms a set of nested groups (clades), in which smaller clades are contained within larger ones. The hierarchic pattern that has become the hallmark of cladistic analysis is related to the lack of transitional forms found between groups. DARWIN expected evolution to leave us with surviving modern groups within groups, but he expected the history of life to proceed in a gradualistic sequence that blurs the lines between groups. The scarcity of such fossil transitions can only be explained in DARWINIAN terms as a sampling problem, an artifact of an incomplete fossil record (DARWIN 2000, p292). Modern paleontologists generally agree, however, that the fossil record is actually robust enough to tell us that the scarcity of transitional forms is real and significant (SIMPSON 1960; GOULD 1977; VALENTINE/ERWIN 1987; DONO-VAN/PAUL 1998; FOOTE 1996; FOOTE/SEPKOSKI 1999), making the hierarchic pattern a genuine aberration in the gradualistic picture.

The priority of typology over continuity has persisted, according to SIMPSON, among "all schools of taxonomy including some that usually oppose typology in principle" (SIMPSON 1961, p49). *Haikouella* contributes to this crystalizing picture of distinct, fully formed body plans from near the start. Developmentalists observe the same hierarchical pro-

Evolution and Cognition | 146 | 2003, Vol. 9, No. 2

cesses at work in both ontogenesis and evolution. Biologist Brian GOODWIN writes: "Developmental processes are hierarchical. So are biological classification schemes" (GOODWIN 1994, p234). Wallace ARTHUR agrees: "A theme running through the work of most contributors to what can now be described as evolutionary developmental biology is the relationship between these two hierarchies", (ARTHUR 1997, p256) and he asserts that "it *is* informative about the nature of evolutionary mechanisms" (p257).

Saltation

How much further back can we trace our ancestors? Nicholas HOLLAND, for one, wants to know what *preceded* these complex, early Cambrian craniates, a question, he says, that remains as big a mystery as ever: "Where are *Haikouella's* ancestors? The sixty-four dollar question is, What is this hooked to? That nobody knows" (HOLLAND personal communication).

In his presentation to an international symposium on Cambrian body plans (1999), HOLLAND gave genetic reasons why the most popular theoretical predecessor for chordates, tunicates (sea squirts), only works in the imagination of the theorists. When chordates are compared genetically with tunicates and fruit flies, he says, "the fruit fly is closer to the tunicate every time" (HOLLAND personal communication).

No obviously ancestral fossils presently exist to support theories about how chordates, or the other phyla, evolved in Precambrian times. "There are a lot of different totally cutup paper doll ideas about where things come from that aren't based on fossils at all, but people sitting in their armchairs", says HOLLAND (personal communication). The ceaseless re-interpretation of ancestral lineages for the phyla is easily demonstrated by the relevant literature (ARTHUR 1997, p73; BERGSTRÖM 1994; LYNCH 1999). Wherever the first chordates came from, HOLLAND thinks science must now take seriously the concept of "saltation", the possibility of evolution in quick jumps. However broadly one defines "saltation", paleontological evidence for the notion is certainly supportive of the internalist/developmentalist position.

Though opinions vary about the Precambrian antiquity of the phyla, all agree that almost all of these most widely separated animal groups had appeared by the early Cambrian period. Why didn't new phyla continue to evolve during subsequent eras? Why did



Figure 3. Sponge embryos seen under the microscope at the cellular level in early cleaving stages, well preserved by the thousands from Precambrian deposit at Weng'an.

such disparate phyla as chordates, mollusks, arthropods, and the 35-or-so others first show up in the fossil record so close to the same time? CHEN places the window of opportunity for the explosive evolution of the majority of body plans within a narrow window of three million years (CHEN 1999), though of course, this is hotly disputed.

Body plans seen in the Precambrian include sponges, annelid worms, and echinoderms (like sea stars), but little else to represent the many lineages expected to lead to the 35 Cambrian groups. Gradualists have claimed that the ancestors of the many other disparate Bauplans must have been too small or too soft to be preserved. But since 1998, phosphate deposits at a Precambrian locale called Weng'an have proved capable of preserving the smallest and softest organisms imaginable (LI et al. 1998). Sponge embryos have been found by the thousands in early cleaving stages, seen under the microscope in groups of 2, 4, 8 cells, etc. (Figure 3). Though small and soft specimens are found in abundance, the number of body plans remains small.

The questions raised by such findings drew sixty scientists to Kunming, China, for a symposium entitled: "The Origins of Animal Body Plans and Their Fossil Records". Perhaps it took the discovery of our own phylum's participation in the early Cambrian big bang to bring together such an international gathering to consider a pattern some call "top-down evolution".

Evolution and Cognition | 147 | 2003, Vol. 9, No. 2

Top-Down Evolution

The appearance of chordates at this early date adds to the evidence for what Berkeley paleobiologist James VALENTINE and his colleagues call a "topdown" pattern in the fossil record (ERWIN/VALEN-TINE/SEPKOWSKI 1987). In the most published diagram in the history of evolutionary biology (and the only diagram in On the Origin of Species), DAR-WIN illustrated what became the standard, bottom-up view of how new taxa evolve (DARWIN 2000, pp514-515). Beginning with small variations, evolving organisms diverge further from the original ancestor, eventually diversifying into new species, then new genera, new families, new orders, and the splitting continues until the highest taxa are reached, which are separated from one another by the greatest differences (DARWIN, p120, p128; SIMPSON 1953, pp383-384).

"The textbooks all teach that evolution takes place when a new species appears, when the morphology is very close", said CHEN in a talk titled "Top-Down Evolution and the Fossil Record" (CHEN 1999). "But that story is not true, according to our fossil finds", he told the assembled scientists. "The new phyla make their start in the early days, instead of coming at the top". He pointed to a very different-looking diagram of his own to illustrate the fact that morphological gaps among animals were greater near the beginning and less significant later (LEWIN 1988; ARTHUR 1977, pp81– 82; SCHWARTZ 1999, p3) (Figure 4).

Rather than observing one body plan branching out into greater numbers of body plans over geologic time, paleontologists instead note maximum disparity¹⁰ between body plans from the beginning, and the retention of essential characters within each throughout geologic history, while increasing diversification occurs at successively lower hierarchical levels; (VALENTINE 1986; PADIAN/CLEMENS 1985; BERGSTRÖM 1994). Developmental geneticist Stuart KAUFFMAN sees deeper reasons for the pattern than anything neo-DARWINISM knows: "the patterns of the branching, dramatic at first, then dwindling to twiddling with details later, are likely to be lawful" (KAUFFMAN 1995, p14).

After listening to CHEN's "top-down" talk, paleontologist David BOTTJER said, "I think the Cambrian explosion is going to tell us something different about evolution, in the sense that it's not the same story that we have always been taught" (BOTTJER personal communication). BOTTJER can't



Figure 4. Origin of the phyla.

argue with the top-down pattern: "After the concentration of phyla first showing up in the Cambrian", he said, "then we see classes, then orders, families, and that's where much of the action is later on, after the Cambrian. So there is that kind of a pattern. And the question is, why is that happening?" Participants in the Kunming symposium came prepared to propose new, sometimes non-DARWINIAN mechanisms to explain the relatively abrupt appearance of the phyla.

New explanations included: saltatory evolution as a reaction to submarine hydrothermal eruptions (YANG et al. 1999); a "Cambrian substrate revolution" in which burrowing animals destroyed the microbial mat habitat of others, resulting in new environments and extensive adaptations (BOTTJER 1999); a billion years of genetic preadaptations for complex metazoans through "set-aside cells" (DAVIDSON 1999); "intelligent design", the inference that the preadaptations and "appearance of design" point to an actual design by an intelligent entity, whether that entity be explained by directed panspermia, a Platonic demiurge, a theistic deity, or some other, unknown intelligent cause (NELSON 1999; WELLS 1999); the evolution of Platonic forms as a vitalistic process, i.e., the suggestion that evolution is driven by a controlling force or principle within organic forms that cannot be reduced to physics and chemistry alone (DENTON 1999); and top-down evolution, in which laws of harmony play at least as great a role in evolution as competition (CHEN 1999).

Contingency

Returning to our original three hypotheses, we now ask: How do findings surrounding the earliestknown craniate affect probabilities for the evolution of cognition? Cephalization prior to the development of an internal body support structure might suggest a body plan in which the head is in some sense dominant. Observing the top-down pattern in the subsequent fossil record, some might further see in this a law-like process dictating an early appearance of brainy chordates among the body plans. But what kind of natural law would demand that, of all the evolving phyla, one of them would necessarily develop a conspicuous brain, ready to be subsequently supported by the vertebrate structure?

Worse, what kind of law would demand that such a pre-backbone craniate would necessarily survive what Stephen Jay GOULD calls "the Burgess decimation"? (GOULD 1989, pp233-239). In Wonderful Life, he suggests that "a 90 percent chance of death would be a good estimate for major Burgess [Cambrian] lineages" (p47). In recent years, Peter WARD and Donald Brownlee have stirred up controversy about the odds against complex life (even as complex as a flatworm) evolving on another planet. In their book Rare Earth, they argue that complex life in the galaxy may be rare, mainly because of the small number of planets that provide enough time and the right conditions for its evolution (WARD/ BROWNLEE 2000). They also believe that the Cambrian explosion of so many new, widely separated, complex animal groups didn't have to happen. Neo-DARWINISM doesn't predict such an event. And the fact that virtually no new animal phyla have evolved in the 530 million years since should give us pause (VALENTINE 1995).

The new discoveries in China take this concern a step further, demonstrating that even a "charmed place" like Earth, apparently ideal for life, is not *necessarily* good enough to produce advanced intelligence. First we learn that chordates, like the other animal phyla, must evolve early to evolve at all (since new phyla don't keep appearing after the Cambrian). Then we learn that major groups did not survive the Cambrian, though we know of no reason why they were less fit than chordates. The first fact (all body plans forming close together in time) has a law-like quality about it, while the second (extinctions) appears highly stochastic.

GOULD may have been overenthusiastic in his use of the term "Cambrian decimation" (GOULD 1989, p47), and we should not infer that chordates only had once chance in ten to survive the Cambrian. To say that most lineages disappeared is not to say that most phyla disappeared. We do not know that the Cambrian ended with a massive extinction event, as we do about the end of five other periods. However, some analyses show that more disappearances occurred by the end of the Cambrian than at the end of any of the "Big Five" extinctions (WARD/BROWN-LEE 2000, p184)—even the Permian, usually declared to be the most catastrophic. According to independent studies by paleontologists Helen Tappan and Norman Newell, about 60 percent of marine families went extinct in the Cambrian, compared to about 55 percent in the Permian" (Ibid).

What we can say with certainty is that craniates had their birth in the most dangerous possible period in the history of metazoan life. As has long been known, in only one period do the number of animal phyla decrease: the Cambrian, and in that period they decrease drastically (DOBZHANSKY et al. 1977, pp421-23). Cambrian researchers say that this period was by far the riskiest because species diversity within each phylum was at an all-time low, making it easier for changing environmental conditions to destroy an entire phylum merely by eliminating a few species (GOULD 2002, p1315). But as geologic time progresses, there is a pattern of increasing diversity at lower taxonomic levels relative to the higher taxa. Today there are far fewer classes and orders than existed four- to five-hundred million years ago, while there are probably eight to ten times the number of species (Dobzhansky et al. 1977, p428).

Thus the same phenomenon that gives rise to the top-down pattern in the fossil record also helps to explain why GOULD considered the chordate's Cambrian survival a momentous event, like winning the lottery. And what reason can we give for expecting our winning streak to hold up through all the subsequent chancy events, including at least five major extinctions? Perfectly fit species were caught by chance at the wrong time, belonging to groups that would not otherwise have gone extinct, but that simply happened to be at a low point in species numbers (since species numbers fluctuate randomly over time) (GOULD 2002, pp1312-1317). The K-T impact that was apparently ultimately responsible for exterminating the dinosaurs 65 million years ago happened to work in favor of small mammals. But what if that extraterrestrial impactor had missed the Earth? Might dinosaurs have ruled the planet for another 200 million years, preventing the evolution of cognition?

Evolution and Cognition | 149 | 2003, Vol. 9, No. 2

The Principle of Mediocrity

Such an idea appears to challenge the Principle of Mediocrity (also known as the Copernican Principle), the assumption that there is nothing special about our place in the universe. After all, the universe does not revolve around Earth. Our planet, our solar system, even our galaxy is but one of billions. Applied to our subject, the Principle of Mediocrity implies that if human-level cognition exists here, it must exist commonly throughout the universe.

What astronomers know by principle and by multiple proofs, biologists are anxious to demonstrate too. Suspecting that we self-aware beings shouldn't be exceptional, biologists and paleontologists are beginning to contemplate new ways to beat the odds. A few even wonder if the game is somehow rigged. This seems to be Jun-Yuan CHEN's position, and a theme of his "top-down" talk at the Kunming conference: the fossil record demonstrates something more than accidental progress by a series of flukes.

Rather than seeing a gradual accumulation of small modifications that finally added up to widely separated animal groups, CHEN observes an explosive appearance of particular forms—sophisticated, widely separated animal groups, right from the start. Diagnostic characters did not accrue over time, but showed up with their first appearance in the form of Bauplans, including our own (CHEN 1999; BERGSTRÖM 1994). To say that this was not in some sense "meant to be" would seem to be a denial of this important, Copernican axiom of science.

Cognition in Other Body Plans?

Haikouella demonstrates that the basic body plan that sets us so far apart from mollusks and arthropods was in place at the beginning of the animal fossil record. Chordates, named for the notochord that would eventually be largely replaced and surrounded by the vertebral column, seem ideally suited to provide the structure required to put sensory organs up high, where they can help an animal get the best perspective on surroundings. Other design requirements for brainy wannabees naturally follow: the brain needs to be near these sensory organs, to minimize reaction time, and the whole should be protected by an encasement. A distinct head is thus a part of the package, which CHEN and SHU claim to have found in these earliest "craniates". But again, the very considerations that make this animal appear to be optimally placed also make its position look tenuous.

Consider a world where chordates had gone extinct with other Cambrian animals. GOULD considers this to be a likelier scenario, a world without fish, birds, reptiles and mammals. Instead, lots of sea stars, crustaceans, insects, and worms. But, we ask, couldn't chordates have re-evolved later? Not when we recall that, with the possible exception of *Bryozoa* ("moss animals"), no new animal phylum has ever evolved since the Cambrian period (VAL-ENTINE 1995). If advanced intelligence was to evolve after that, it would have had to take a radically different form.

In that case, wouldn't another animal group have filled our niche to eventually develop the ability to compose literature and do math? Again, not likely. Biologists have reasons to doubt that other phyla are so well suited to developing large brains situated in a commanding position. For a simple thought experiment, readers should try to picture a sea star, bug or worm with a big head. Or, more to the point, readers might try to think of a member of a non-chordate phylum on this planet that did develop a written language and technology, given 500 million years to do so.

Paleobiologist Michael BENTON points out that "the vertebrate design lends itself to the development and protection of a brain. This organ is present in other animals, but there are limits on its growth-one of them imposed very early in the history of life, when animals were first developing basic equipment like a front and a back, sense organs, and the ability to use information from the sense organs ..." (BENTON 1993). BENTON notes the importance of the right architecture to create space available for the cluster of nervous tissue where data arrive and orders depart. While vertebrates separate this central ganglion from the rest of the body, arthropods and mollusks wrap it around their gut. Observes BENTON: "Any tendency for this tissue to grow is likely to squeeze the tube of the gut and constrict the supply of food. This is a contradiction that the arthropod design has never resolved..." (Ibid).

What if chordates survived, but not mammals or primates? Some might argue that, given more time, dinosaurs themselves could have developed high intelligence. Paleobiologists, however, say that a wholly different kind of skull would be required. "You cannot simply grow a giant brain in a dinosaur like Velociraptor: you have to reconstruct the skull", writes Richard FORTEY. "Consciousness is not a clever trick to be whipped up from any set of neurons like a soufflé from an egg" (FORTEY 1998).

Evolution and Cognition | **150** | 2003, Vol. 9, No. 2

Partly because our present existence appears to depend upon a long string of unpredictable accidents, biologists know of no fundamental "law of progress" to show them why the path should have led to anything like Homo sapiens. Biologist C. O. LOVEJOY writes that "the evolution of cognition is the product of a variety of influences and preadaptive capacities, the absence of any one of which would have completely negated the process" (LOVEJOY 1981). He notes that the human's complex nervous system is actually a reproductive liability, requiring a longer gestation period and a longer time to train the young. LOVEJOY concludes: "It is evident that the evolution of cognition is neither the result of an evolutionary trend nor an event of even the lowest calculable probability, but rather the result of a series of highly specific evolutionary events whose ultimate cause is traceable to selection for unrelated factors such as locomotion and diet" (Ibid).

"If intelligence has such high value", writes Ernst MAYR, "why don't we see more species develop it?" (MAYR 1996). He contrasts the singular development of high intelligence with the repeated evolution of sight, which occurred at least 40 times (SALVINI-PLAWEN/MAYR 1977). He calls the search for extraterrestrial intelligence "hopeless" and "a waste of time", concluding that "for all practical purposes, man is alone" (MAYR 2001, p263).

The list of leading biologists and paleontologists on record for defending this intelligence-by-fluke position is impressive, including SIMPSON, DOBZHANSKY, FRANCOIS, AYALA, and GOULD (BAR-ROW/TIPLER 1986, p133). British astronomer John BARROW and American physicist Frank TIPLER note that "there has developed a general consensus among evolutionists that the evolution of intelligent life, comparable in information-processing ability to that of *Homo sapiens*, is so improbable that it is unlikely to have occurred on any other planet in the entire visible universe" (Ibid).

Many astronomers who once took optimistic positions on the probability of finding signals from an extraterrestrial intelligence are adjusting their predictions. Forty years of null SETI results may have even taken their toll on optimist Robert JAS-TROW, director of the Mt. Wilson Observatory. Though he once told this writer, "We'll be hearing from those guys soon", he has since modified his statement to *"If life is common,* we'll be hearing from those guys soon" (JASTROW personal communication). Even this guarded claim shows an astronomer's willingness to believe that the route from life to intelligence is an obvious one, which, as we have seen, is disputed by most biologists and paleontologists schooled in the Modern Synthesis.

Something Missing

Some paleontologists see such strong trends in the fossil record that they don't believe contingent events can overcome them. Jun-Yuan CHEN believes that there must be other forces driving evolution toward intelligence besides natural selection and mutations. If evolution were restricted to these two forces, he says, then all life would still be microbial. "Bacteria are very successful", pointed out CHEN. "They have a great capacity to adapt to environmental changes" (personal communication). And he noted that bacteria have flourished better than other life forms that have come and gone over billions of years without complexity or intelligence. Complex life, CHEN said, is less capable of making adaptations, so that "complex, highly evolved life, like the human, has no reason to appear. So why should these chance mutations plan such complex types of animals?" (Ibid). What's missing from neo-DARWINISM?

Wallace ARTHUR pictures neo-DARWINIAN theory as a grand edifice with foundations and walls that are composed of interdependent disciplines, so that "if one part turns out to be wrong, the whole structure may eventually collapse" (ARTHUR 1997, p285). Until the developmental component has made its contribution, he says, "There is not just a brick or two missing, but rather a whole section of the building" (Ibid).

Physicist Paul DAVIES suspects that biologists have concluded too rashly that they understand life's origin and evolution, and that "we are missing something very fundamental about the whole business" (DAVIES 1999, p17). Cosmologists routinely use the term "anthropic principle" to describe the many preconditions for complex life met by severely constricted universal constants (BARROW/TIPLER 1986; BARROW 2002; GREENE 1999). These include the apparent "fine-tuning" of the universe's expansion rate (sometimes calculated to be "tuned" to one part in 10⁶⁰ at one second after the big bang, as a precondition for life) (HAWKING 1988, pp121–122; KRAUSS 1998) and the precise strengths of nature's four fundamental forces (e.g., the strength of the electromagnetic force appears to be tuned relative to the gravitational force to at least one part in 10³⁶, as a precondition for the existence of stable stars) (BAR-ROW/TIPLER 1986, p219; DAVIES 1983, p188; REES

Evolution and Cognition | **151** | 2003, Vol. 9, No. 2

1999, p2). DAVIES has long wondered if biologists would see the constraints and the bio-friendly pattern too.

CONWAY MORRIS sees something like it: "Consider, for example, the sponges", he writes, "which by general consent are the most primitive living metazoans. Nevertheless, their biochemistry includes elements that seem to foreshadow the immune system of vertebrates" (CONWAY MORRIS 2000; SCHÄCKE et al. 1994). Though sponges do not have nerve cells, they already have neuronal-like receptors, so that they "seem to be almost 'animals in waiting'" (CONWAY MORRIS 2000). CONWAY MORRIS believes that caution is in order and that such findings can be carried too far, producing a distorted view; yet he continues listing examples of what appear as preadaptations, such as the nervous system of amphioxus revealing "a vertebrate in waiting" (Ibid).

Similarly, recent genetic studies of hemichordates, which have no brains, show that these most plausible models for proximate ancestors to chordates already contain the genes that express the brain and spinal cord in vertebrates (LACALLI 2003). Hemichordate genes that are responsible for patterning the body along its front-to-back axis were found expressed in the surface tissue in a nearly identical arrangement to those that express themselves in vertebrate brains and spinal chords (LOWE et al. 2003). LOWE et al. favor the idea that a complex genetic map was in place long before the complex morphology.

The bottom line, according to CHEN, is that the standard mechanisms of neo-DARWINISM offer no basis for a "ladder of progress". So far, a noncontroversial view. But if his "top-down" alternative gains acceptance, it would create a paradigm shift in biology. His replacement of competition with harmony and top-down evolution could be taken to suggest the first rungs in such a guiding ladder. CHEN's discovery of Haikouella shows that the last really big turn in the pathway to humanity did not occur at the end of the evolutionary process, but at the beginning. Does this mean that the "goal" of humanity was set from the beginning of metazoan life? Few other participants at the Kunming conference were willing to say anything like that. But some did, including New Zealand geneticist Michael DENTON.

Arguing from the fact that almost no new phyla evolved after the Cambrian explosion, DENTON said: "The body plans of the Cambrian are probably built into nature from the beginning" (DENTON 1999). DENTON is part of a team that recently revealed how, at its base, life follows "laws of form" in the discrete,

three-dimensional folding patterns of protein molecules. The folds can be classified into a finite number of structural families that are determined by natural law, not natural selection-much like the physical laws that give rise to atomic elements in the periodic table. Writing for the Journal of Theoretical Biology, his team describes the protein folds as "'lawful forms' in the Platonic and pre-DARWINIAN sense of the word, which are bound to occur everywhere in the universe where the same 20 amino acids are used for their construction" (DENTON/MAR-SHALL/LEGGE 2002). In another piece, for Nature, DENTON and MARSHALL argue: "If forms as complex as the protein folds are intrinsic features of nature, might some of the higher architecture of life also be determined by physical law?" (DENTON/MARSHALL 2001).

Moreover, given the limitations of a material world of flux, DENTON considers the possibility that "the laws of nature are fit for only one unique thinking being capable of acquiring knowledge and ultimately comprehending the cosmos" (DENTON 1998). He cites Mark WARD's research on the fine balance achieved (1) between the size/number of neurons and the blood vessels which nourish them, and (2) between the width of axons and the required insulation/blood supply (WARD 1997). Referring to this and to the staggering compaction of synaptic connections in the human brain, he writes that "the evidence is certainly consistent with the possibility that the human brain does indeed represent the most advanced information-processing device that can be built according to biological principles" (DENTON 1998).

However, to say that the experience of consciousness is fully explained by the physical laws that produce such a brain is a non sequitur, except to committed reductionists. Physicists from Brian PIPPARD to Stephen WEINBERG have raised questions about the reasonableness of expecting consciousness itself¹¹ to ever be subsumed under the domain of physics and chemistry (PIPPARD 1992; WEINBERG 1992, p44). Given a complex structure with ample computing power, should a theoretical physicist be able to deduce the existence of self-awareness from laws of physics? Cognative scientist David J. CHALM-ERS suggests that the problem of trying to derive consciousness from physical laws is so troublesome that any final theory of physics "must contain an additional fundamental component". He proposes "that conscious experience be considered a fundamental feature, irreducible to anything more basic" (CHALM-ERS 1995).

Evolution and Cognition | 152 | 2003, Vol. 9, No. 2

Concluding Options

If nature is somehow rigged in favor of mind, then the tremendous odds against our existence disappear. But if that

concept were to catch hold in scientific circles, Paul DAVIES claims that it would create a "decisive shift" in science (DAVIES 1999, p263), reversing a 300-year trend toward reductionist thinking. We cannot at the same time hold to the Principle of Mediocrity and to the idea that human cognition is a bizarre case.

The evidence surrounding the discovery of the earliest craniates forces us to choose between renouncing one of two deeply embedded traditions of modern science. Either mind plays a role in nature by necessity, which appears to contradict the reductionist basis for doing science-or mind plays no role and has appeared as an "oddball rarity", which contradicts science's equally cherished Copernican Principle. This means that our first two original options-human-level cognition as either an accidental, or a law-like, process—will give us serious problems either way we choose. If we choose the lawful process option, we must then ask ourselves: What kind of law will ensure that primates (or any other form preadapted for braininess) will survive through the bottleneck of contingent events that are beyond the control of any known natural mechanisms?

To opt for human-level cognition as both accidental rarity *and* commonplace occurrence is to render both options meaningless, since they contradict each other. We do have a third option: that our existence is primarily due to neither accident nor cosmic law. To speak awkwardly, as we did at the beginning of this article, of the human-level cognition "observed" on Earth is to flagrantly ignore our own unique position as both observer and the observed. The inside information we're privy to as conscious and frequently conscientious primates may provide some hints about the workings of chance and natural law, for our lives would seem to be, from our own viewpoints, composed of more than either accidents or laws. From an unlikely combination of cir-

Author's address

Fredric J. Heeren, 903 E. Park Street, Olathe, KS 66061, USA. Email: hee@flash.net

cumstances have emerged beings who are much more than the sum of their parts. It would seem that our most uniquely human abilities are not predictable in any detail

from our morphologies.

If we say that we transcend our physical world with our human achievements-our music, literature, humor, love-it still remains for us to decide whether this transcendence emerged by accident or according to a prior purpose. Simon CONWAY MOR-RIS suggests that this may be the principal reason that biologists have hesitated so long to explore directionality and channeling: "If evolution is in some sense channeled, then this reopens the controversial prospect of a teleology; that is, the process is underpinned by a purpose". (CONWAY MORRIS 1998, p14). And he notes a growing trend to bring cosmology's Anthropic Principle down to our biosphere. CONWAY MORRIS sees humanity's uniqueness in our ability to make these kinds of choicesand voices his irritation with those who choose to live irresponsibly based on an assumption of life's purposelessness (Ibid). The reductionist's belief in human life as a cosmic accident is a metaphysical commitment too.

After all, at least to this point, the most dazzling thing on Earth that evolution has done is to produce volitional beings whose present lives have little to do with the physical processes that brought them. "Uniquely", CONWAY MORRIS writes, "there is inherent in our human situation the possibility of transcendence" (Ibid). The fact that it's only a possibility speaks volumes, once again, about the human capacity to choose.

Acknowledgements

I wish to thank the referees for providing many useful suggestions for improvement of this manuscript. I also thank Jun-Yuan CHEN, Nanjing Institute of Geology and Palaeontology, and Paul K. CHIEN, University of San Francisco, for generously contributing photographic materials and diagrams.

Notes

- 1 Biological dictionaries now frequently replace the subphylum name *Vertebrata* by the newer, broader phylum name *Craniata* in order "to represent the distinguishing characteristics more accurately" (RUDIN 1997). Chief among craniate distinctions is a manifest head containing a brain and sensory organs. Modern craniates are also characterized, as vertebrates were, by a segmented vertebral column. The group continues to include fish, amphibians, reptiles, birds, and mammals. In modern cladograms, the *Chordata* clade includes the *Craniata* clade, the *Craniata* clade includes the *Vertebrata* and *Myxinoidea* clades, etc.
- 2 The following distinction is made here between *cognition* and *intelligence: Cognition* is used to describe the application of mental processes involved in knowledge; while *intelligence* describes the *ability* to know, regardless of its use. *Cognition* is the act of using of one's *intelligence*. Thus the human capacity for thought and reason called *intelligence* results in human-level *cognition*, an awareness involving reasoning and judgment apparently unlike the mental processes of any other animal on this planet. Human-level cognition should be detectable, since it tends to find expression in human-level communication, engineering feats, abstract and mathematical problem-solving, musical compositions, fine art, literature, science, etc.
- **3** That is, the evolution of human-level cognition is dependent upon a long series of unpredictable, historical events, making its occurrence on Earth a rarity. If other planets harbor life, only a very tiny fraction, if any, would then be expected to host human-level-or-higher cognition.
- **4** One might argue for a fourth option: that human-level cognition exists as both a rare fluke *and* a common or lawlike property of the universe; but the statements can both be true only by rendering them meaningless. While there is nothing logically contradictory about chance mutations (flukes) and natural selection (law) working together to produce novel forms of life, the question here is whether it is rare or common for any such combination of law and chance to produce forms that result in *human-level cognition*. When referring to the evolution of cognition, the first two hypotheses are contradictory and do not allow for both as a primary cause.
- **5** The terms *body plan* and *Bauplan* are generally used interchangeably. James VALENTINE applies the term Bauplan to "the upper levels of the taxonomic hierarchy" where "phyla- or class-level clades are characterized by their possession of particular assemblages of homologous architectural and structural features" (VALENTINE 1986). Wallace ARTHUR identifies six morphological characters to distinguish animal body plans: skeleton, symmetry, pairs of appendages, body cavity, cleavage pattern, and segmentation (ARTHUR 1997, p27). Like others, ARTHUR tends to identify animal body plans in the Cambrian period with the animal phyla (he speaks of the Cambrian "origin of the 35 or so animal body plans" (ARTHUR 1997, opening page), though in more

general contexts (non-Cambrian) he speaks of "phylum/ class level body plans" (ARTHUR 1997, p27).

6 *Developmentalism:* emphasizes the importance of understanding ontogeny—the history of, and the genetic processes involved in, the development of the individual organism—for understanding evolution.

Neo-DARWINISM: emphasizes natural selection and mutations as the overwhelming driving forces for understanding evolution. Also called the Modern Synthesis (since it synthesizes these two mechanisms).

Formalism: emphasizes internal constraints toward the evolution of particular body forms.

Functionalism: emphasizes external adaptations as the primary force behind the production of characters that function best in particular environments.

Punctuationalism: emphasizes the geologically abrupt origin and subsequent stasis ("equilibrium") of most species. *Gradualism:* emphasizes the slow and constant accretion of small changes that eventually add up to larger changes and separations between organisms.

Top-down theory: emphasizes the evolution of the higher taxa first, so that the most widely separated groups appear early, and "the diversification of the phyla occurs before that of classes, classes before that of orders, orders before that of families" (ERWIN/VALENTIN/SEPKOWSKI 1987).

Bottom-up theory: emphasizes the evolution of the higher taxa from the accumulation of lower taxa, creating a phylogenetic tree of increasing diversity and eventual disparity.

- **7** Reductionism is a philosophical method of explaining a complex set of facts by reducing them to a set of smaller, simpler facts; the whole should be predictable from its smaller, constituent parts.
- **8** *Constraints* may be negative or positive; negatively, they are restrictions on evolution's direction; positively, they are preferred directionality of variation; either internal or external factors may *constrain* evolution toward particular forms. *Channels* are usually positive, internal, preferred evolutionary pathways.
- **9** *Convergence* is the explanation for shared characters of independently evolved organisms. In GOULD's lexicon, the convergence of characters is based upon common external adaptations. He carefully distinguishes convergence from *parallelism*, which is the independent origin of common features channeled by *internal* constraints of homologous genes or developmental pathways. Other scientists frequently employ the term *convergence* to include any case where the evolution of characters repeats itself, whether explained by external constraints or internal channeling.
- **10** *Disparity* is the word usually used to describe differences between organisms that involve whole body plans; *diversity* is reserved for differences between lower-taxa organisms, especially at the species level (GOULD 1989, p49).
- **11** WEINBERG distinguishes between "consciousness itself", the self-awareness/feelings experienced by humans, and "correlatives to consciousness" that may be examined in terms of brain waves, electrical activity, hormones in the blood, etc. (WEINBERG 1992, p44).

References

- Arthur, W. (1988) A Theory of the Evolution of Development. Wiley: Chichester UK.
- Arthur, W. (1997) The Origin of Animal Body Plans. Cambridge University Press: Cambridge UK.
- Barrow, J.D. (2002) The Constants of Nature. Pantheon Books: New York.
- Barrow, J. D./Tipler, F. J. (1986) The Anthropic Cosmological Principle. Oxford University Press: Oxford/New York.
- Benton, M. (1993) The rise of the fishes. In: Gould, S. J. (ed) The Book of Life. W. W. Norton & Company: New York, pp. 65–77.
- **Bergström, J. (1994)** Ideas on early animal evolution. In: Bengston, S. (ed) Early Life on Earth, Nobel Symposium no. 84. Columbia University Press: New York, pp. 460–466.
- **Bottjer, D. J. (1999)** The Cambrian substrate revolution and the origin and evolution of the mollusca. Presented at the International Symposium of the Origin of Animal Body Plans and their Fossil Records, Kunming, China, June 20–25, 1999 (Sponsored by the Early Life Research Center and the Chinese Academy of Sciences).
- **Chalmers, D. J. (1995)** The puzzle of conscious experience. Scientific American 273: 80–86.
- Chen, J.-Y. (1999) Top-down evolution and the fossil evidence. Presented at the International Symposium of the Origin of Animal Body Plans and their Fossil Records, Kunming, China, June 20–25, 1999.
- Chen, J.-Y., Dzik, J./Edgecombe, G. D./Ramsköld, L./Zhou, G.-Q. (1995) A possible Early Cambrian chordate. Nature 377: 720–722.
- Chen, J.-Y./Huang, D.-Y./Li, C.-W. (1999a) An early Cambrian craniate-like chordate. Nature 402: 518–522.
- Chen, J.-Y./Huang, D.-Y./Li, C.-W. (1999b) An early Cambrian new chordate casting new light on the origin of craniate. Presented at the International Symposium of the Origin of Animal Body Plans and their Fossil Records, Kunming, China, June 20–25, 1999.
- **Conway Morris, S. (1998)** The Crucible of Creation: The Burgess Shale and the rise of animals. Oxford University Press: Oxford/New York.
- **Conway Morris, S. (1998–1999)** The challenge: Showdown on the Burgess Shale. Natural History 107: 48–55.
- **Conway Morris, S. (2000)** Evolution: Bringing molecules into the fold. Cell. 100: 1–11.
- **Conway Morris, S./Whittington, H. B. (1979)** The animals of the Burgess Shale. Scientific American 240: 131.
- **Darwin, C. (2000)** On the Origin of Species by Means of Natural Selection. Harvard University Press: Cambridge MA. Originally published in 1859.
- Davidson, E. (1999) Regulatory control of embryological development and the evolutionary origin of animals. Presented at the International Symposium of the Origin of Animal Body Plans and their Fossil Records, Kunming, China, June 20–25, 1999.
- Davies, P. (1983) God and the New Physics. Simon and Schuster: New York.
- Davies, P. (1999) The Fifth Miracle. Simon & Schuster: New York.
- **De Beer, G. R. (1930)** Embryology and Evolution. Oxford University Press: Oxford UK.
- **Dennett, D. C. (1996)** Darwin's Dangerous Idea: Evolution and the Meaning of Life. Penguin Books: London.
- Denton, M. J. (1998) Nature's Destiny. The Free Press: New York.
- Denton, M. J. (1999) The nature and evolution of biological

form. Presented at the International Symposium of the Origin of Animal Body Plans and their Fossil Records, Kunming, China, June 20–25, 1999.

- Denton, M. J./Marshall, C. J. (2001) Laws of form revisited. Nature 410: 417.
- **Denton, M. J./Marshall, C. J./Legge, M. (2002)** The protein folds as platonic forms: New support for the pre-Darwinian conception of evolution by natural law. Journal of Theoretical Biology. 219(3): 325–342.
- Dobzhansky, T./Ayala, F./Ledyard Stebbins, G./Valentine, J. (1977) Evolution. W.H. Freeman & Co: San Francisco.
- Donovan, S. K./Paul, C. R. C. (eds) (1998) The Adequacy of the Fossil Record. Wiley: Chichester UK.
- Dzik, J. (1995) Yunnanozoon and the ancestry of chordates. Acta Paleontologica Polonica, 40: 341–360.
- Enserink, M. (1999) Fossil opens window on early animal history. Science 286: 1829.
- **Erwin, D. H./Valentine, J. W./Sepkowski, J. J. (1987)** A comparative study of diversification events: the early Paleozoic versus the Mesozoic. Evolution 41: 1183.
- Foote, M. (1996) On the probability of ancestors in the fossil record. Paleobiology 22(2): 141–151.
- Foote, M./Sepkoski, J. J. Jr. (1999) Absolute measures of the completeness of the fossil record. Nature 398: 415–417.
- Fortey, R. (1998) Shock lobsters. Book review of "The Crucible of Creation: The Burgess Shale and the Rise of Animals" by S. Conway Morris. London Review of Books 20(19).
- **Goodwin, B. (1994)** How the Leopard changed Its Spots: The Evolution of Complexity. Weidenfeld & Nicholson: London.
- Gould, S. J. (1977) Evolution's erratic pace. Natural History 86: 12–16.
- Gould, S. J. (1989) Wonderful Life: The Burgess Shale and the Nature of History. W. W. Norton & Company: New York/ London.
- **Gould, S. J. (1996–1997)** War of the worldviews. Natural History 105: 22–33.
- **Gould, S. J. (1997)** Darwinian fundamentalism. The New York Review of Books 44(10): 2–4.
- **Gould, S. J. (2002)** The Structure of Evolutionary Theory. Harvard University Press: Cambridge MA.
- Greene, B. (1999) The Elegant Universe: Superstrings, Hidden Dimensions, and the Quest for the Ultimate Theory. W. W. Norton & Company: New York.
- Hall, B. G. (1992) Evolutionary Developmental Biology. Chapman & Hall: London.
- Hawking, S. W. (1988) A Brief History of Time: From the Big Bang to Black Holes. Bantam: New York.
- Holland, N. D. (1999) Bringing the chordate ancestor of the vertebrates into focus: A decade of progress in neontology and paleontology. Presented at the International Symposium of the Origin of Animal Body Plans and their Fossil Records, Kunming, China, June 20–25, 1999.
- **Kauffman, S. A. (1995)** At Home in the Universe: The Search for Laws of Self-Organization and Complexity. Viking-Penguin: London.
- Krauss, L. M. (1998) The end of the age problem, and the case for a cosmological constant revisited. Astrophysical Journal 501: 465.
- Lacalli, T. (2003) Body plans and simple brains. Nature 424: 263–264.
- Lewin, R. (1988) A lopsided look at evolution. Science 242: 292.
- Li, C.-W. (1999) In: Chiu, Y.-T. Fossil discovery changes theory. The Taipei Times, December 2, 1999.
- Li, C.-W./Chen, J.-Y./Tsu, E. H. (1998) Precambrian sponges
- Evolution and Cognition | **155** | 2003, Vol. 9, No. 2

with cellular structures. Science 279: 879-882.

- **Lovejoy, C. O. (1981)** Evolution of man and its implications for general principles of the evolution of intelligent life. In: Billingham, J. (ed) Life in the Universe. MIT Press: Cambridge MA, pp. 317–329.
- Lowe, C. J./Wu, M./Salic, A./Evans, L./Lander, E./Stange-Thomann, N./Gruber, C. E./Gerhart, J./Kirschner, M. (2003) Anteroposterior patterning in hemichordates and the origins of the chordate nervous system. Cell 133: 853–865.
- Lynch, M. (1999) The age and relationships of the major animal phyla. Evolution 53(2): 319–324.
- Mayr, E. (1996) The search for extraterrestrial intelligence: Scientific quest or hopeful folly? A debate between Ernst Mayer and Carl Sagan. Planetary Report 9(3): 4–14.

Mayr, E. (2001) What Evolution Is. Basic Books: New York.

- Monastersky, R. (1999) Waking up to the dawn of vertebrates. Science News 156: 292.
- **Nelson, P. A. (1999)** Generative entrenchment and body plans. Presented at the International Symposium of the Origin of Animal Body Plans and their Fossil Records, Kunming, China, June 20–25, 1999.
- Padian, K./Clemens, W. A. (1985) Terrestrial vertebrate diversity: Episodes and insights. In: Valentine, J. W. (ed) Phanerozoic Diversity Patterns: Profiles in Macroevolution. Princeton University Press: Princeton, pp. 41–96.
- **Pippard, B. (1992)** Review of "Understanding the present: Science and the soul of modern man", by B. Appleyard. Nature 357: 29.
- **Purnell, M. A./Donoghue, P. C. J. (1997)** Architecture and functional morphology of the skeletal apparatus of ozarkodinid conodonts. Philosophical Transactions of the Royal Society of London, Series B 352: 1545–1564.
- Raff, R. A./Kauffman, T. C. (1983) Embryos, Genes and Evolution: The Developmental-Genetic Basis of Evolutionary Change. Macmillan: New York.
- Rainey, P. (2003) Review of "Evolution: Five big questions", by M. Russell/A. Pomiankowski/G. Turner/P. Rainey/R. Dunbar. New Scientist, June 14, p. 32.
- **Rees, M. (1999)** Just Six Numbers: The Deep Forces that Shape the Universe. Basic Books: New York.
- Rudin, N. (1997) Dictionary of Modern Biology. Barron's Educational Series: New York.
- Salvini-Plawen, L./Mayr, E. (1977) On the evolution of photoreceptors and eyes. In: Hecht, M. K./Steere, W. C./Wallance, B. (eds) Evolutionary Biology. Plenum: New York, pp. 207–263.
- Schlosser, G. (2000) Amphibian variations: The role of modules in mosaic evolution. Presented at the 5th Altenberg Workshop on Theoretical Biology 2000, Konrad Lorenz Institut, Altenberg, Austria.
- Schäcke, H./Rinkevich, B./Gamulin, V./Müller, I. M./Müller, W. E. G. (1994) Immunoglobin-like domain is present in the extracellular part of the receptor tyrosine kinase from

the marine sponge Geodia cydonium. Journal of Molecular Recognition 7: 273–276.

- Schmalhausen, I. I. (1949) Factors of Evolution: The Theory of Stabilizing Selection. Blakiston: Philadelphia PA.
- Schwartz, J. H. (1999) Sudden Origins: Fossils, Genes, and the Emergence of Species. John Wiley: New York.
- Shu, D./Conway Morris, S./Zhang, X.-L. (1996) A pikaia-like chordate from the lower Cambrian of China. Nature 384: 157–158.
- Shu, D./Luo, H.-L./Conway Morris, S./Zhang, X.-L./Hu, S.-X./ Chen, L./Han, J./Zhu, M./Li, Y./Chen, L.-Z. (1999) Lower Cambrian vertebrates from south China. Nature 402:42– 46.
- Shu, D./Zhang, X./Cheng, L. (1996) Reinterpretation of Yunnanozoon as the earliest known hemichordate. Nature 380:428–430.
- Simpson, G. G. (1953) The Major Features of Evolution. Simon and Schuster: New York.
- Simpson, G. G. (1960) The history of life. In: Tax, S. (ed) Evolution after Darwin. Volume 1: The Evolution of Life. University of Chicago Press: Chicago, pp. 117–180.
- Simpson, G. G. (1961) Principles of Animal Taxonomy. Columbia University Press: New York.
- Solé, R./Goodwin, B. (2000) Signs of Life: How Complexity Pervades Biology. Basic Books: New York.
- Valentine, J. W. (1986) Fossil record of the origin of baupläne and its implications. In: Raup, D. M./Jablonski, D. (eds) Patterns and processes in the history of life. Springer-Verlag: Berlin, pp. 209–222.
- Valentine, J. W. (1995) Why no new phyla after the Cambrian? Genome and ecospace hypotheses revisited. Palaios 10(2): 190–194.
- Valentine, J. W./Erwin D. H. (1987) Interpreting great developmental experiments: The fossil record. In: Raff, R. A./ Raff, E. C. (eds) Development as an Evolutionary Process. Alan R. Liss: New York, pp. 71–107.
- Waddington, C. H. (1957) The Strategy of the Genes. Allen & Unwin: London.
- Ward, M. (1997) End of the road for brain evolution. New Scientist 153(2066): 14.
- Ward, P. D./Brownlee, D. (2000) Rare Earth. Springer-Verlag: New York.
- Weinberg, S. (1992) Dreams of a Final Theory. Pantheon: New York.
- Wells, J. (1999) Resolving the paradox of homeotic genes in the origin of animal body plans. Presented at the International Symposium of the Origin of Animal Body Plans and their Fossil Records, Kunming, China, June 20–25, 1999.
- Yang, W./Chen, N./Wang, X./Gao, L./Zhao, Y./Guo, Q. (1999) Saltatory evolution mechanism of the Weng'an biota. Presented at the International Symposium of the Origin of Animal Body Plans and their Fossil Records, Kunming, China, June 20–25, 1999.

Gestalt Experiments and Inductive Observations

Konrad Lorenz's early Epistemological Writings and the Methods of Classical Ethology

 $E_{\rm some\ crucial\ insights}^{
m thology}$ brought and perspectives to the study of behavior, in particular the idea that behavior can be studied within a comparativeevolutionary framework by means of homologizing components of behavioral patterns and by causal analysis of behavior components and their integration. Early ethology is well-known for its extensive use of qualitative observations of animals under their natural conditions. These observations are combined with experiments that try

Abstract

During the 1940s Konrad LORENZ formulated his early epistemological views, focusing on the cognitive mechanisms induction and Gestalt perception. After the war he used this philosophical framework to defend the approach of classical ethology against other approaches to animal behavior. The present paper examines the relationship between LORENZ's ethological methodology and his philosophy of science and knowledge. The main aim of LORENZ's post-war epistemological writings is to provide an epistemological and cognitive theory of observation in order to defend the observational approach of classical ethology against the view that it was no rigorous science.

Key words

Konrad LORENZ, Gestalt perception, observation, ethology.

to analyze behavioral patterns and establish specific claims about animal behavior. Nowadays, there is still disagreement about the significance of observation and experiments and their relation. ALLEN (forthcoming) points out that in debates about the interpretation of animal cognition the interpretations of those ethologists who have actually spent time watching the animals are favored by some biologists. In fact, some practitioners of cognitive ethology feel that experiments at best just confirm what one already knows, while others are more skeptical about interpretations of observations without experiments.

As a major representative of ethology Konrad LORENZ not only shaped its approach and made the methods and theories of ethology known to the biological community and the interested public, he also embedded his defense of the ethological and comparative approach in a philosophical theory of scientific knowledge. Gestalt perception and induction are the main eleof LORENZ'S ments epistemology. These ideas were originally formulated during the second world war (see for instance the Russian Manuscript, LORENZ 1948). In the post-war period LORENZ emphasized his epistemological ideas as part of an attempt to justify his ethological approach and his views about the role of observation and experiment. The aim of the present paper is to call attention to LORENZ'S philosophical

account and how it integrates with his biological methodology.

Observation and Experiment

LORENZ's account of the general character of observation and experiment remains unchanged throughout his writings. The specific relationship between these two aspects is exhibited by LORENZ's early biological work as well as his later explicit methodological account of them (LORENZ 1935; LORENZ/TINBERGEN 1938; LORENZ 1948, 1981). LORENZ emphasizes four distinctive features of classical observational practice: observations are to be carried out in a hypothesis-free way; the approach consists in doing qualitative observations rather than quantitative measurements; animals are to be observed in their natural environment; and observations include all features of the organisms and

Evolution and Cognition | 157 | 2003, Vol. 9, No. 2

their environment and rely on a large observational basis. One should not think that ethology (in particular the more developed ethology of the post-war period) always proceeded in this fashion. Instead, LORENZ's bold claims about how ethology is done are based on some of his opinions as to how science should be done or how a science should develop (see BURKHARDT 1981). In any case, the way classical ethology was done was noticeably different from other contemporary approaches to animal behavior such as comparative psychology or later on sociobiology. In fact, LORENZ encountered methodological critique from both groups.

LORENZ defends the ethological approach of gathering observational data without having a hypothesis in mind against the view that scientific accounts must be based on a theory (1948, p216; unless otherwise indicated, references refer to a publication of LORENZ). LORENZ thinks that it is not only possible to make observations devoid of a theoretical framework, he also points to the possibility of confirmation bias as a disadvantage of basing one's observational approach on a hypothesis (1948, pp31, 71). Instead of forming theoretical considerations at an early stage of inquiry, scientific inquiry has to start with "observation pure and simple" (1958, p246). LORENZ calls this "presuppositionless observation" (1981, p47) or often "unbiased observation" (1958, p250; see also 1948, p213; 1950 p131; 1959, p281). Now, even amateur naturalists who dedicated all their time to field observations of animals did not necessarily refrain from interpreting their observations and relate them to biological theories. Edmund Selous, for instance, used his extensive observations to justify theories of sexual selection. In the case of LORENZ, the theoretical assumption that behavioral characters can be used to characterize taxonomic groups and reconstruct phylogenies guided his observational practice from very early on. But LORENZ is right insofar as before 1935 he and TINBER-GEN completely lacked an interpretative framework, unlike most people working in animal psychology. After the war ethology became more theoretical, but compared to other approaches such as comparative psychology or sociobiology ethology was not that much an enterprise driven by explicitly formulated theories and hypothesis that were put to empirical test. In particular sociobiology took quantitative hypotheses based on models from population genetics as the starting point and used observations primarily to test these hypotheses.

In addition, early ethology relied on qualitative observations of behavior. Instead of making quanti-

tative measurements and recording them, observations are recorded using written descriptions (1935, pp112ff), drawings and photographs (LORENZ/TIN-BERGEN 1938). This qualitative approach of classical ethology has to be defended against an understanding that *only* allows for quantitative measurement and statistical evaluation as a means of scientific objectivity (1958, pp246, 256; 1959, p281; 1963a, p1; 1981, pp40ff, 68ff). Despite LORENZ's emphasis on qualitative observation, post-war ethology did very well make use of quantitative measurements, but connected them with qualitative observations (SCHLEIDT/SCHLEIDT 1958 is an example).

A further feature of the ethological way of observing is to systematically study animals in their natural ecological environment. Animals are kept "in an environment as close as possible to their natural habitats, for the purpose of general biological and specifically ethological observations" (1935, pp108-109). Keeping animals under laboratory conditions suffers from the drawback that the organisms do not exhibit their natural adaptive behavior patterns and thus functional and artifactual behavior are likely to be confused. This approach of observing animals under natural conditions continues the "amateurism" of the forerunners of classical ethology (1981, p47). However, for LORENZ the optimal method is not to study animals in the field. For in this case, it is difficult and extremely time-consuming to be able to observe the relevant behavior patterns (1948, p222). The chapter "Animal Keeping as a Research Method" of the Russian Manuscript gives a detailed account of how the study of animal behavior has to be organized (1948, pp221ff). Several animals of a species have to be kept so that they live in proximity to the researchers and in a controllable environment. Apart from conducting observations in that way, animal keeping makes it possible to conduct experiments under otherwise natural conditions (1948, p222). LORENZ's practice of raising and keeping animals was quite similar to the practice of his mentor Oskar HEINROTH or the American zoologist Charles Otis WHITMAN. WHITMAN's and LORENZ's practice of animal keeping enabled them to simultaneously observe the behavior of several closely related species, which is hard to achieve in the field. But LORENZ's approach differed from the early field naturalist in the British tradition (Selous and Howard) or in the Dutch tradition (RÖELL 2000). In particular Niko TINBERGEN heavily relied on field observations (and experiments) instead of animal keeping (TINBERGEN 1932, 1935, 1951), while LORENZ viewed field work as a control for observa-

Evolution and Cognition | 158 | 2003, Vol. 9, No. 2

tions gained from animals kept in semi-natural conditions (1981, pp47ff). Thus TINBERGEN represented the ecological dimension of ethology more fully than LORENZ did.¹ Despite the fact that LORENZ was not a field naturalist, his approach was quite distinct from the American tradition of animal psychology and animal behavior studies. LORENZ often labeled the American scientists 'behaviorists', despite the fact that scientists such as John WATSON or Robert YERKES had quite different views of animal minds and behavior. But apart from theoretical differences, the American tradition was definitely characterized by laboratory studies. Observations and experiments under laboratory or controlled conditions dominated (WATSON 1914; YERKES 1925), field studies were an exception (e.g., WATSON 1908 or YERKES' student NISSEN 1931). This is due to the fact that in the United Studies animal behavior was studied at universities and institutes from early on, while some of the forerunners of European ethology were amateur naturalists. (In the United States, there were a few important field naturalist among the zoologists, rather than the psychologists. But even zoologists primarily studied behavior in laboratories; DEWSBURY 1988.)

Ethological observations do not simply focus on specific aspects of animal behavior. Instead, according to LORENZ's ideal ethological analysis start with considering all features of an individual including its natural environment (1981, pp52f). Observation has to give an overview of the complete behavioral system of a species (1948, pp221f). Only in this manner it is possible to obtain knowledge about the relevant substructures of the system, its functions and relations to other organisms and the environment (1950, p120). LORENZ calls this approach that tries to take all features into account an "analysis on a broad front". It is justified by the fact that the object under study is a complex entity whose parts interact mutually. The chapter "The Organism as an Entity and Analysis on a Broad Front" of the Russian Manuscript (1948, pp137ff) gives a discussion and defense of this method (see also 1958, p249). To achieve this general overview of the whole behavioral complex of a species it is necessary to have a large observational basis that often may require years to obtain (1935, p109; 1948, p222; 1981, p48). Even before LORENZ, the naturalist Edmund Selous claimed that in his field notes he really wrote down everything he saw (BURKHARDT, forthcoming). While this is literally impossible, in particular the early practitioners in ethology are characterized by spending most of their time observing animals. Charles Otis WHITMAN, for instance, kept detailed notes of his extensive observations, and carried out observations for years before he published his results (LILLIE 1911). Like LORENZ, he emphasized that a detailed and extensive knowledge of the behavioral repertoire of a species is crucial to avoid false interpretations and conclusions (WHITMAN 1899a, 1899b).

The role of experiments within ethology is nicely illustrated by LORENZ's and TINBERGEN's classical study of the egg-rolling behavior of the Greylag goose (1938). (TINBERGEN was in fact the person who was more skilled in devising and conducting experiments. Compare TINBERGEN 1932, 1935, TINBERGEN/ KUENEN 1939 with LORENZ's more observational and theoretical papers LORENZ 1932, 1935, 1937.) On LORENZ's account, experiments are to be preceded not only temporally, but also methodologically and logically by observations that are not goal-directed. For an experiment only makes sense if the natural units and their interaction are known to a sufficient degree (1981, p53). An encompassing observational basis is necessary to have a grasp of the structure and function of the system under consideration. Knowledge about the structure of organisms and the function of behavior is needed for having a biologically meaningful classification of the parts of the system studied. These elements identified by observation can one after the other analyzed in more detail and substantiated by experiments designed for addressing such particular questions (1935, pp110f). Thus, in the ethological approach, the experiment is necessarily posterior to observation.

LORENZ is for instance impressed by the detailed observational work of the ornithologist Margaret M. NICE, and he states that one should always have such a detailed knowledge of an animal before one starts to conduct experiments.

"Ich war begeistert von der Forschungsweise dieser Dame. *So* müßte man eben *jedes* Vieh kennen, *bevor* man mit ihm 'Versuche' anzustellen beginnt." (LORENZ, letter to STRESEMANN, Blatt 94–97).²

The purpose of experiments is to analyze the subcomponents of an overall behavioral pattern (1935, pp105, 233; 1937, p293). The experiment enables one to keep the constitutive elements of behavior apart, and to study a largely dissociated component in detail. A certain experimental design allows the researcher to answer specific questions about the function of behavioral component and the causal relationship between the components. Dummy experiments serve to determine the specific stimulus that triggers an innate releasing mechanism and

Evolution and Cognition | **159** | 2003, Vol. 9, No. 2

thus help to determine whether this stimulus is learned or innate (1935, p143, p228; LORENZ/TINBER-GEN 1938, p333). Other experiments figure out whether a stimulus or drive is of external or internal origin (1950, pp135f). A type of experiment that was very important for classical ethology is the deprivation experiment. It consists of depriving animals of certain environmental stimuli during their development. This is to show that behavioral components which nonetheless are exhibited after the deprivation period are not learned but must be innate (1965, pp83ff; 1981, pp57ff). LORENZ's mentor HEINROTH made extensive use of this technique by hand-rearing birds in relative isolation to discern their instinctive behavior features (HEINROTH/HEIN-ROTH 1924-1933).

Induction and Gestalt Perception

Induction, on LORENZ's account, is a rational cognitive process, which means for LORENZ that inductive reasoning is a conscious process (1948, p30). More precisely, a rational cognitive process is characterized by the fact that the individual is aware of the assumptions made and in particular can verify the steps of reasoning. Throughout LORENZ's writings, induction is conceptually clearly contrasted with non-rational processes such as intuition and Gestalt perception (1948, p55; 1950, p166; 1954, p198; 1959, pp302, 306; 1981, pp43f). Rational cognitive processes are in principle open to introspection and the individual has control over the way of reasoning, which contrasts them to subconscious processes (1959, pp283, 312f; 1981, p41). A general feature of induction is that it abstracts regularities and principles from knowledge about single facts (1954, p198). The distinctive feature of inductive science is that it aims at proving assertions by collected evidence (1948, p54). The last word on verifying the correctness of a scientific hypothesis has to be achieved by quantitative analysis (1958, p246).

The advantage of induction stems from the fact that it is a rational cognitive process. The inductive base is explicitly known, the single steps in inductive reasoning can be analyzed and verified. In particular the way of reasoning that led to the conclusion can be communicated to other persons in a manner that allows for critical assessment of the scientific claim. The bigger inductive basis, the better the scientific result is substantiated (1948, p64). The main disadvantage of induction consists in the difficulty of anticipating empirically important or theoretically fundamental results. Induction provides good evidence for the results obtained from data, but it is unlikely that induction arrives at scientifically relevant conclusions that might be substantiated by an adequate inductive basis (1948, p65).

A central view of LORENZ on empirical science and research is the idea that inductive research proceeds in three stages. The terms used to describe this categorization (the idiographic, the systematic and the nomothetic stage) are adapted from the philosopher Wilhelm WINDELBAND (1894). An early detailed account of this philosophy of science is given in the chapter "Induction" of the Russian Manuscript (1948, pp28ff) and it is maintained throughout LORENZ'S writings (1958, p251; 1959, p283; 1963a, p2). The first phase of inductive science is the idiographic stage. It consist of gathering data, observations, and descriptions in a loose manner. The aim is to obtain an inductive basis that is as large as possible. The systematic stage is the next step. It proceeds by analyzing the data, comparing the various bits of evidence, finding interrelations among the single items, and categorizing them. In the first two stages, the inductive basis gets assembled in a hypothesis-free and theory-independent fashion (1950, p129). The final phase is the nomothetic stage, which consist in formulating laws and general principles supported by the evidence. It is characterized by a process of abstraction (1948, p30). Whereas the idiographic and the systematic stage cannot by kept clearly apart, the nomothetic phase can be more sharply separated from the foregoing stages (1948, p29). Observation, description, comparison, and systematization are "indispensable steps that must have been gone through before the first attempt at the abstraction of natural laws, of nomothesis in WINDELBAND's classical term, is undertaken" (1958, p262). The reliability of the results obtained by induction is proportional to the broadness and the scope of the inductive basis (1948, p31; 1950, p194; 1959, pp302, 315). LORENZ claims that early ethology developed just in this three stage fashion (1950, p131; 1948). However, one needs to keep in mind that such statements are due to the fact that LORENZ thought that any real science has to emerge in this fashion. BURKHARDT (1981) makes clear that some of LORENZ's historical statement should not be viewed as an attempt to reconstruct the history of ethology, but they stem from LORENZ's attempt to create ethology as a discipline, including a defense of ethology as a discipline that developed the way a real science ought to develop on LORENZ's view.³

Evolution and Cognition | 160 | 2003, Vol. 9, No. 2

A peculiar feature of LORENZ's epistemology is Gestalt perception, a cognitive process that is of paramount importance for the post-war defense of his approach. The notion of Gestalt perception is present in LORENZ's early ethological writings, but as an epistemological idea explicitly formulated for the first time in early philosophical works such as the Russian Manuscript and defended throughout LORENZ's philosophical, methodological, and biological writings (1943, 1948, 1950, 1958, 1959, 1963a, 1963b, 1977, 1981). The most elaborated account is given in the article Gestalt perception as a source of scientific knowledge (1959). In Germany and Austria Gestalt psychology was an influential school before the second world war (ASH 1998), and so LORENZ could make use of ideas available from the German-speaking psychological tradition. In fact, as a student LORENZ took classes and interacted with the Vienna psychologist Karl BÜHLER and his assistant Egon BRUNSWIK, both of whom studied the psychology of perception. On LORENZ account, the distinctive feature of Gestalt perception is that it is a ratiomorphic process. LORENZ borrows this term in the post-war period from BRUNSWIK (1952, 1955), but the idea was present in BRUNSWIK's post-war work even though he did not use this particular term for it (see BRUNSWIK 1934). Before using BRUN-SWIK's term LORENZ makes reference to Hermann von HELMHOLTZ's concept of unconscious inference (1948, p57; see HELMHOLTZ 1925, vol. 3.) Talking about ratiomorphic mechanisms means that we are dealing with a non-rational, subconscious process, that nonetheless exhibits strong analogies to rational thought-in particular induction-with respect to how it operates (1948, p55; 1959, pp296, 302; 1981, p41). For LORENZ Gestalt perception as a special kind of perception is obviously a neurophysiological mechanism. LORENZ uses form constancy as an example to illustrate a simple type of Gestalt perception (1959, pp302f). When an observer sees an object from different sides (e.g., because the object moves and turns), the object is recognized as the same entity with the same form even though the image on the retina changes continuously. The human perceptual apparatus is able to extract from the different sensory data the relevant information that makes one see the same object at different instances. This process obviously works unconsciously. The individual is not aware of the steps taken by the nervous system to process information; only the result (e.g., the recognition of the object) is mediated to consciousness (1959, p296; 1981, p43). Nonetheless, this subconscious mechanism exhibits analogies to reasoning (1959, p302; 1981, pp42f). The process starts by using incoming data as 'evidential basis', it draws 'inferences' from this data and comes to a 'conclusion'. As this procedure must make certain implicit 'assumptions' as to how objects in the external world behave, this kind of perception can be fooled by experimentally creating a situation that is unlikely to occur under normal conditions and violates these 'assumptions' (1959, pp297ff; 1981, p41). In the preceding description of a ratiomorphic process the use of terms refering to rational thought in scare quotes indicates the analogy with rational cognitive processes. The ratiomorphic cognitive apparatus—which is evolved-has built-in 'assumptions' and 'hypotheses' about how the external world looks like. As LORENZ assumes that Gestalt perception is important for hypothesis-free observation, these built-in implicit 'hypotheses' need to be kept apart from the hypothesis that a scientists consciously entertains. Both in rational and ratiomorphic mechanisms the reliability of the inference is proportional to the broadness of the inductive basis (1959, p315). LORENZ repeatedly mentions that Gestalt perception is analogous to rational thought in that it makes inferences and conclusions, albeit in an unconscious manner (1948, pp57, 59; 1958, p253; 1959, p283). His discussion of constancy mechanisms shows that the kind of inference that is made is similar to rational abstraction, which is an important part of induction (see also 1958, p252). Form constancy is a mechanism of pattern recognition; from various bits of incoming stimuli the features enabling the recognition of objects are mediated. At a few places LORENZ explicitly states that Gestalt perception is analogous to rational abstraction (1951, p173; 1958, p252; 1959, pp283, 304).

Form constancy is a simple example of Gestalt perception, but there are more complicated types of it. Gestalt perception is able to extract similarities and regularities out of data sets consisting of miscellaneous items. For instance, it is able to recognize objects and individuals given by perceptual information. Moreover, it can detect natural kinds and categories that exists in nature. By means of Gestalt perception different individuals are perceived to belong to the same species, genus, or family (1959, pp306f). Gestalt perception is an important tool in systematics. This is of fundamental importance for LORENZ because he views biological features as taxonomic characters that can be homologized. Ethology as a comparative-phylogenetic approach to behavior can makes fruitful use of Gestalt perception.

Evolution and Cognition | 161 | 2003, Vol. 9, No. 2

In Gestalt perception "many elements are always combined into one entity" (1948, p58). This is an issue where the idea of the quality of a Gestalt is stressed by LORENZ. On his account a Gestalt quality is a type, something beyond concrete individuals (1959, pp306ff). Finally, this cognitive process can even be used to discover empirical principles and law-like relations out of a large amount of information containing otherwise irrelevant data (1948, p64; 1959, pp282f, 310). This makes Gestalt perception a powerful cognitive mechanism. In short, Gestalt perception detects natural units and unexpected principles and lawful regularities. For this reason, in the Russian Manuscript this latter complex function of Gestalt perception is called intuition (1948, pp30, 54).

Gestalt perception is claimed to be important for all three stages of empirical research, as it seems because Gestalt perception takes in the data from the first two stages (operating independent of hypotheses) and creates the abstractions and general principles that are characteristic for the third, the nomothetic stage (1948, p30; 1963a, p7). In addition, LORENZ says that not only research based on rational induction proceeds according to the three stage model, but that Gestalt perception itself implicitly proceeds in this fashion (1959, p283).

Comparing LORENZ's views on Gestalt perception with the theory of the early Egon BRUNSWIK, by whom LORENZ was probably influenced, reveals some commonalities.⁴ BRUNSWIK emphasizes the fact that perception is analogous to reasoning (1934, pp2, 50, 127), but in contrast to reasoning it is immediate, i.e., only the result of the perceptive process is mediated to consciousness, but intermediate steps are not (p1). Like LORENZ, BRUNSWIK assumes that perception is directed at gaining knowledge about the objective properties of objects (1934, p.V; BRUNSWIK 1937), which was an important issue for Karl BÜHLER's general approach, too. Even though perception can be fooled under experimental conditions, BRUNSWIK states that under normal conditions perception is very reliable. In fact, he compares perception with instinct/conditioning. Instincts and conditioned behavior operate reliably in standard cases, but in contrast to insight they are inflexible so that they cannot adapt do new situations. Perception has a performance that is similar to instincts or conditioned behavior (1934, pp114ff). While LORENZ might agree with this analogy he does not make use of it, but contrasts ratiomorphic processes-which are usually reliable but cannot adapt to new situations-just with rational

processes. BRUNSWIK assumes that there is continuum between measurement and perception, so that measurement is just a limiting case of perception (1934, p9). LORENZ, in contrast, seems to make a sharper distinction between rational and ratiomorphic processes.

The cognitive capacities that are important for LORENZ are induction and Gestalt perception. In this sense, LORENZ's epistemology includes nothing but these two features. This is clear from the Russian Manuscript (1948) and from Gestalt perception... (1959). LORENZ'S more general evolutionary epistemology (see for instance LORENZ 1977) is wellknown, so one needs to give some remarks about the relationship between evolution and the already discussed cognitive processes. In fact, LORENZ formulated some of his ideas about evolutionary epistemology before giving his first detailed account of induction and Gestalt perception in the Russian Manuscript (see LORENZ 1941, 1943). For LORENZ it is clear that our perceptive apparatus, including Gestalt perception, and its capacities to generate reliable knowledge is the product of evolution (1958, pp252f; 1959, p289; 1963a, p6). However, LORENZ does not give an elaborated account of the evolution of the human ratiomorphic apparatus. Gestalt perception is discussed in section 7.2 of Behind the Mirror (1977), but the discussion is very short (6 pages) and does not go much beyond LORENZ's usual remarks. Gestalt perception is not only evolved, but there are learned aspects of it (1951, pp166ff; 1977, p216). Gestalt perception needs training and it is differently developed in different individuals (1959, p313; 1973, p8). So LORENZ could have given an account of how the innate and learned aspects of Gestalt perception integrate. In addition, not only ratiomorphic, but also rational cognitive processes are evolved and part of LORENZ's evolutionary epistemology. But LORENZ does not give a detailed account of how rational and ratiomorphic processes integrate and how the distinction between these different types of cognitive processes came about in the course of evolution. Independent of evolutionary issues, in the post-war period LORENZ stresses the impotence of Gestalt perception as a cognitive mechanism. However, he does not offer a development and elaboration of his ideas and does not refer to the psychological literature (in particular new results) on perception. Rather, LORENZ repeats what are in his view the basic features and faculties of Gestalt perception and sometimes mentions a few classical ideas of Gestalt psychologists.

Evolution and Cognition | 162 | 2003, Vol. 9, No. 2

Mapping Gestalt Perception/Induction onto Observation/Experiment

The first issue to be discussed is the question of in what manner the epistemological apparatus (ratiomorphic Gestalt perception and rational induction) relates to the ethological methodology (observation and experiment) as LORENZ viewed it. Some of LORENZ's writings (especially the *Russian Manuscript*) suggest an epistemic asymmetry between Gestalt perception and induction and attribute different roles in scientific research to these mechanisms, so that a straightforward Gestalt perception–observation and induction–experiment correspondence seems plausible.

The different epistemic roles of induction and Gestalt perception are due to the fact that the former is a rational process, but the latter is not. Induction is a tool for scientific analysis and for the confirmation of scientific claims. It is a cognitive process that aims at objective knowledge. The Gestalt, on the other hand, is a "purely subjective phenomenon" (1948, p140).⁵ The result of this perceptive cognitive process is mediated as a whole to subjective experience. Without its parts being open to analysis, the Gestalt is seen and the received entity accepted (1948, p58). This epistemic difference is in particular shown by the fact that a result obtained by Gestalt perception is refuted by inductive evidence contradicting that result (1948, p68). The strength of induction is its accountability, while Gestalt perception is characterized by its incorrigibility and unaccountability (1948, p64). For this reason, Gestalt perception and induction play a different role in scientific research. Whereas induction is shortsighted and therefore not a good tool for scientific discovery (1948, p65), Gestalt perception is able to see unexpected regularities and unforeseen lawfulness (1948, p63; 1959, p282; 1981, p44). It is a good "hunch generator" (1963a, p7). As only Gestalt perception is really able to fulfill this function, its role in scientific research is to lead the way of discovery (1948, p56, 64; 1963a, p8). Gestalt perception is only a means of discovery. A detected principle has to be substantiated by induction, it is the job of rational processes to confirm scientific claims.

"Therefore, only induction can increase the reliability of a result and only induction can validate a result that has been obtained exclusively through intuition!" (1948, p65)

This yields the following account. *The (subjective) discovery of scientific principles is achieved by Gestalt perception, while the (objective) confirmation and justi-*

fication is obtained through induction. So far the identification of this epistemic relationship between Gestalt perception and induction in LORENZ's work has mainly been based on the Russian Manuscript. Whereas these two different roles are very explicit in this piece of work, some passages of the later article *Gestalt perception...* give the same picture (1959, p316).

On the one hand, LORENZ three stage model of science sounds like a sort of naïve Baconianism, according to which one first needs to collect in an unbiased and theory-independent manner as much data as possible and only then one can generalize laws and theories from it. On the other hand, the idea of Gestalt perception proposing hypothesis to be rationally tested might incline one to assume that LORENZ has a hypothetico-deductive account of science. Neither is really the case. LORENZ definitely is not a Popperian falsificationist. On his account, hypothesis cannot only be disconfirmed, but also confirmed by evidence. LORENZ accepts induction, in fact, he states that one can "use the breadth of the inductive basis to assess with genuine mathematical accuracy a probability value for the correctness of the result. Where that breadth is sufficient, the probability is so close to certainty that we can confidently equate the two." (1948, p64) This makes him sound somewhat like a proponent of the modern Bayesian approach to theory confirmation by evidence. However, Bayesianism is just about confirmation and does not tell us how to come up with theories to be verified. The same applies for a naïve Baconian approach, which cannot make sure that anything theoretically useful emerges from unbiased and unfocussed data gathering. The discovery of theories was somewhat neglected by traditional philosophy of science, but LORENZ addresses this issue by pointing to Gestalt perception, which is supposed to deliver unexpected principles and hypotheses. This is maybe the most fruitful aspect of LORENZ's epistemological perspective (besides the fact that he makes use of an evolutionary framework). LORENZ proposes processes that generate theories as well as confirm them. Rational and ratiomorphic processes are considered psychologically founded mechanisms, which in combination address the two most important aspects about scientific rationality-discovery and justification of theories.

As discussed in the preceding section, experiment are methodologically and logically posterior to observation. Useful experiments can only be conducted *after* a good deal of observation has been undertaken. In fact, observation gives a *meaningful*

Evolution and Cognition | 163 | 2003, Vol. 9, No. 2

account of the structure and function of the observed system and experiments can only then verify details about the components of the behavior of a species and analyze their causal interaction. In an analogous manner, Gestalt perception has to be employed first to detect interesting principles that are to be confirmed by induction. This suggests that the two step procedure Gestalt perception-induction can be mapped onto the observation-experiment procedure in a manner such that both correspond to each other. This interpretation means that observation in the ethological approach is largely driven by the cognitive process of Gestalt perception, while mainly the cognitive function of induction is used in the experimental phase. The fact that the experiments fulfill the analytic demands of inductive science became clear in my exposition of this method. As will be discussed in more detail in the next section, Gestalt perception actually corresponds in several respects to observation. Indeed, a main function of LORENZ's account of the importance of Gestalt perception is to defend his observational approach.

However, both in the Russian Manuscript and in Gestalt perception... there are passages that do not fit the simple interpretation given so far. The relation between Gestalt perception and induction is more complex. First, according to the references given above Gestalt perception and induction have an epistemically different status insofar as only the latter provides objective knowledge and in the case of conflict the result obtained by a rational process is to be preferred. Nonetheless, Gestalt perception and induction are not completely different with respect to their objectivity. In particular when defending the qualitative approach of ethology against views that consider only quantitative measurements as objective, LORENZ's points out that all cognitive processes are in a sense subjective (1959, p320). In addition, perception tells us about the properties of objects in the external reality (1959, p301). In fact, it is the only source of knowledge about the reality surrounding us (1981, p41). Under standard condition-given that the 'evidential basis' is right-results obtained by Gestalt perception are true (1948, p62). Thus Gestalt perception is reliable, despite the fact that its results need to be rationally verified by induction.

Second, in *Gestalt perception*... LORENZ explicitly states that the ratiomorphic mechanism of Gestalt perception and the rational process of induction cannot be sharply separated and that their functions (discovery/verification) intergrade. "It is quite definitely a simplification of this kind to represent the interaction between the various cognitive processes, as I have done above, as if there were always a distinct separation between the prior discovery of an inherent principle through ratiomorphic processes and its subsequent verification through rational processes." (1959, p320)

Quantification, which belongs to inductive analysis, is claimed to be dependent on Gestalt perception (1958, p256; 1959, p320). In fact, Gestalt perception is a precondition for rational reasoning at all (1948, p30). On the other hand, "rational, quantifying, statistical and surveying pre-treatment is necessary to permit Gestalt-formation" (1959, p320). There is a necessary cooperation of Gestalt perception and analysis/experiment (1981, p54). Rational and ratiomorphic processes are strongly entangled and the different steps of inductive science need Gestalt perception (1948, pp28ff). In the Russian Manuscript there are passages that sometimes suggest a strict separation and at other places an interdependency of both types of processes. Gestalt perception... proceeds by first suggesting a clear distinction between rational and ratiomorphic functions, which is then explicitly relativized (compare pp304–319 with 319–322).

Thus LORENZ states that the relation between rational and ratiomorphic processes is complex and that both processes need to be highly entangled in the effective generation of knowledge. Nevertheless, he does not explain sufficiently how he views the interplay between these two processes in detail. (In Gestalt perception... he basically devotes two pages to this issues; see 1959, pp320f). The difficulty is that LORENZ tries to combine rational and ratiomorphic processes (which are of a different nature on his account) without a clear account of how they combine in practice. Despite these complexities, on my interpretation LORENZ still sees a parallel between the ethological methodology (observation and experiment) and the philosophical epistemology (Gestalt perception and induction). Not only are Gestalt perception and induction strongly entangled, observation and experiment also intergrade. Both are needed for ethology and in practice there is an interplay between them. Observation usually precedes experiment and suggests relevant experimental questions. Correspondingly, Gestalt perception leads the way of scientific investigation by proposing hypotheses that need to be confirmed by means of induction. LORENZ makes clear that Gestalt perception is crucial for observation, whereas experiments fulfill the role of scientific induction.

Evolution and Cognition | 164 | 2003, Vol. 9, No. 2

Lorenz's Justificatory Efforts

A good deal of LORENZ's writings are devoted to the defense of his views on biology and scientific knowledge. The next question with respect to the relationship of observation, experiment, Gestalt perception, and induction is how these items justify each other. For instance, does LORENZ's position on how observations are to be made imply the need for Gestalt perception as a cognitive mechanism used in ethological research, or rather are philosophical remarks on Gestalt perception intended to justify the specific ethological observational approach? In LORENZ's writings two groups of ideas can be identified that function as justificatory primitive principles. This means that they are ideas which play an important role in justifying other items of the framework, but which are not themselves justified by other ideas. A group of primitive principles is largely taken for granted or justified internally by claims from the same group of ideas. (Figure 1 gives a rough illustration of the justificatory relationships existing between the different items of LORENZ's theoretical account.)

The first group of ideas are LORENZ's views on induction, which have already been discussed. These are his general ideas on science and rational reasoning, including rational induction as a cognitive mechanism and in particular LORENZ's views of how empirical science works, i.e., the distinction between the idiographic, systematic, and nomothetic stages of science. (There are obviously other views of science that do not assume a theory-free idiographic stage.) The second complex of primitive principles-which despite its paramount importance for LORENZ has not yet been discussed—are related to LORENZ's understanding of "Ganzheit", often translated as *entirety* or (systemic) entity (see 1948, pp137ff; 1950, pp120f; 1958, pp248f; 1959, pp281f; 1981, pp36ff). The main ideas in this complex are claimed to stem from the recognition that biology deals with complex wholes. A systemic entity or an entirety is a system that has subsystems that mutually interact which each other. Every part depends on the other parts. For LORENZ this implies that it is indispensable to study the complex system as a whole. In the case of biological entities this means that one has to take both the structure and the functions of the entirety and its subcomponents into consideration. Research on organisms includes the study of their overall behavior and their environment. LORENZ calls this approach 'analysis on a broad front'.6

Entirety → Gestalt perception

LORENZ's uses this second complex—the ideas about entirety-to justify the need for Gestalt perception as a cognitive tool. Gestalt perception is ideal for studying complex systems and the nervous system/ behavior are especially complex (1959, p283; 1981, p46). One of the main steps in dealing with an organic entirety is to get an overview of its parts. The best way to do so is to make use of Gestalt perception (1981, p47). The entirety approach makes it necessary to study the function of a system in its context as well, and Gestalt perception is a good tool for this. In addition, Gestalt perception is a cognitive capacity that is ideal for a comparative approach. As already outlined, it is able to detect generic kinds and natural units (1948, pp60f; 1959, pp306ff; 1981, pp45f). In general, Gestalt psychology was a very useful approach for LORENZ's perspective. It was a tradition that-for the most partaimed at a rigorous and scientific causal-explanatory framework, in accordance with LORENZ emphasis on causal-analytical thinking in science as opposed to psychological-teleological speculations (1942). But Gestalt psychology was not just reductionistic and instead provided an objective account of holistic features inherent in the perception of the Gestalt as a quality. The perception of some objects is different from the collection of individual perceptions; the Gestalt is a unity. Thus Gestalt psychology allowed LORENZ to maintain his rigorous-scientific attitude and emphasize the fact that biological objects also need to be understood and studied as complex wholes (see 1948, 1951).

Entirety → observation

The ideas about entirety also justify observation as LORENZ understands this method. The recognition of an individual as an organismic entity that has a functional context and a history implies the need for studying healthy animals in their natural environment (1948, pp213, 221, 223; 1981, p40). Another point constitutive of LORENZ's observational approach is the fact that the researcher has to be familiar with all details of the behavioral patterns of an organism. Again, the systemic entity approach makes this obvious (1948, p215; 1981, p38). In other words, recognizing the demands of an analysis on a broad front in the case of organisms and their behavior means that observations have to be performed (at least in some aspects) in the manner of classical ethology.

Induction → gestalt perception

The preceding discussion has already pointed to the fact that LORENZ's views on induction justify the use of Gestalt perception. Induction has the function of confirming scientific claims. But as it is short-sighted, Gestalt perception as a mechanism that finds unforeseen lawfulness is needed to carry out successful inductive science. Gestalt perception leads induction the way (1948, pp56, 64). LORENZ's account of how the rational and ratiomorphic processes operate makes clear that induction can hardly do without Gestalt perception.

Induction → observation

The claims about induction justify directly the approach to observation. LORENZ emphasizes that the nomothetic stage can only be reached after the hypothesis-free idiographic phase (1948, p216; 1950, p129). The crucial function of the idiographic stage is to provide an inductive basis that is as broad as possible including all relevant features (1948, p213; 1950, p131). New explanatory principles can only be abstracted from a sufficiently broad base (1958, p250). Furthermore, the premature formation of a hypothesis can distort the inductive basis. When a researcher has a specific theory in mind it is likely that the data of the idiographic stage are biasedly evaluated so that the scientist sees his or her own hypothesis supported (1948, pp31, 71). This gives a justification for the idea that observation has to provide a large body of data, that it includes all features of the organism and its environment, and in particular that it has to be conducted in hypothesis-free way (1959, p281).

Induction → experiment

Induction also implies the need for experiments. An important part of scientific induction is to provide an account of the details of the object under study, to give a causal analysis, as well as to confirm hypotheses. As has already been discussed the main function of ethological experiments is to establish claims about the specific components of behavior and to examine their (causal) relationship.

Experiment → Gestalt perception

I argued above that the inductive approach justifies the use of Gestalt perception because successful induction presupposes Gestalt perception. The analogous relation that experiment is subordinated to observation might suggest that the conduct of experiments justifies the need for observations. Some of LORENZ's remarks to the effect that experiments do not make sense without observation seem to amount to a direct justification of observation by experiment. However, it hardly amounts to a justification of the specific characteristics of LORENZ's observational approach. I prefer the interpretation that the need for experiments mainly justifies the use of Gestalt perception, which in turn calls for observation (as we will se below). In other words, the necessity of experiments justifies the specific aspects of the observational approach only indirectly, namely insofar as Gestalt perception is the tool for conducting observations. The use of Gestalt perception follows from the need of doing experiments because experiments are useful only if the substructures and functions of a system are known to a certain degree (1981, pp53, 65). Gestalt perception is the best cognitive mechanism to tackle this question (1981, p47). It can break down a complex system into meaningful parts on which an experiment can focus.

Gestalt perception → observation

The following quotations illustrate how self-evident it is for Konrad LORENZ that observation and Gestalt perception go together:

"... in the observation of complex animal behaviour patterns, one can literally see the same process thousands of times without noticing the inherent principle until—quite abruptly—on the following occasion the Gestalt is distinguished from the background of accidental features ..." (1959, p306)

"At the other extreme are the died-in-the-wool behaviourists who deny that Gestalt perception and thus observation of organisms in their natural environment—has any value or even scientific character." (1959, p319)

In fact, the use of Gestalt perception actually justifies the ethological approach to observation. First, the effectiveness of Gestalt perception in finding important regularities and principles is proportional to the amount of data of which Gestalt perception can make use (1948, p63; 1959, p305; 1981, p46). An advantage of this ratiomorphic process is its retentive memory and the large amount of data it can take into account (1959, pp309, 314f; 1981, p44). Gestalt formation can suddenly occur after a long period of data collection (1959, p306; 1981, p45). This lends support to the habit of observing animals over a long span of time (1959, p315; 1981,

Evolution and Cognition | 166 | 2003, Vol. 9, No. 2

p45). In addition, Gestalt perception deals with the whole (1948, p139; 1959, p306; 1981, p46). It is able to get an overview of a multitude of aspects of a system and can extract a lawful relationship out of it (1948, p64). For this reason, effective use of Gestalt perception can be made if all behavioral patterns of organisms are observed in their natural context. As a perceptual mechanism Gestalt perception has qualitative information as input. This justifies the qualitative approach to observation (1948, p64). Finally, Gestalt perception does not need a hypothesis (1948, p63). It works best when the researcher is relaxed and contemplates his object of study while unconsciously collecting data (1959, p316; 1981, pp45f). In fact, rational reasoning about details of the object negatively influenced the operation of Gestalt perception (1959, pp314, 317). In this case, the features of Gestalt perception show that observation has to be conducted without rational influences based on theories or hypothesis. To sum up, LORENZ tries to make clear that the distinctive aspects of the ethological observational approach is chosen in accordance with a cognitive enterprise that relies on Gestalt perception. Observations yield effectively knowledge because this powerful ratiomorphic mechanism is employed. The idea that Gestalt perception is predominantly used for ethological observations (as well as for systematics) can be found throughout LORENZ's writings (1948, pp61f, 213ff; 1954, pp197ff; 1958, pp251ff, 278; 1959, pp306ff, 311, 315; 1963a, pp8f; 1981, pp40ff).

Conclusion

I have argued that on LORENZ's view the observation-experiment relation is largely analogous to the Gestalt perception-induction relation. Experiment needs foregoing observation so that it can be conducted meaningfully. Similarly, induction needs Gestalt perception that leads induction the way. Observation and Gestalt perception discover new principles, whereas experiment and induction have to give an analytic confirmation. Gestalt perception is the crucial cognitive tool for ethological observation, while rational induction is important for experimental analysis and confirmation. LORENZ makes clear that the relation between Gestalt perception and induction is very complex. However, he does not give a concrete account of the interplay between these two cognitive mechanisms and their relation to the biological approach. For this reason, while the Gestalt perception-induction relation can be largely mapped onto the observation-experi-



Figure 1: Overview of LORENZ's justificatory system.

ment relation, it is not clear in detail what the actual role of rational processes in observation is and to which extent Gestalt perception is also important for ethological experiments.

The result of my reconstruction of LORENZ's justificatory relationships between the ideas of entirety, Gestalt perception, induction, observation and experiment are illustrated in Figure 1.7 LORENZ's work exhibits a justificatory system in which all of his relevant ideas are logically related. What becomes clear is the fact that the ideas about entirety/systemic entity (Ganzheit) and the claims about induction and inductive science function as primitive principles. They are not justified by other items, but both are used to justify the need for Gestalt perception as a cognitive tool and the necessity of observations in an ethological manner. Furthermore, three different items lend support to the use of Gestalt perception as a cognitive tool: the views about entirety, induction, and experiments. It is not very surprising that Gestalt perception is justified from different perspectives. While this cognitive mechanism is very important for LORENZ's epistemology, it is hardly endorsed by other post-war biologists. On LORENZ's account Gestalt perception has very strong capacities, e.g., finding scientific hypotheses. LORENZ needs to justify the use and need of this non-rational (albeit ratiomorphic) process that he uses to stress as a tool for scientific inquiry.

Finally, my analysis shows that the primary aim of LORENZ's justificatory efforts is the ethological approach to observation. Even though LORENZ associates Gestalt perception with observation, it is not the case that the specific features about the observational approach are used to justify the use of Gestalt perception. Rather, the ideas about Gestalt perception justify the way observations were carried out by

Evolution and Cognition | 167 | 2003, Vol. 9, No. 2

Ingo Brigandt

the founders of ethology. Indeed, LORENZ's remarks about Gestalt perception are intended to show that there is an important and powerful cognitive mechanism that is able to get knowledge out of what has been observed. The discussion shows that there exist several items that are used to justify the observational approach, but the tenets about observation are not employed to justify any of the other items. Induction and Gestalt perception justify the hypothesis-independency of observation, and the qualitativeness of the approach is substantiated by Gestalt perception. The ideas about entirety justify the fact that organisms are to be studied in their natural environment, while the view that all features have to be observed is supported by all three items (induction, Gestalt perception, and entirety).

From a historical point of view it is highly plausible that the observational method is the main target of Konrad LORENZ's justificatory efforts. For one of the main intentions of LORENZ's post-war defense of his works was to show that non-quantitative, largely theory-independent observations are in fact scientific (this is explicit in 1958, pp246, 256; 1959, p281; 1963a, p1; 1973, p1; 1981, pp40, 68ff). LORENZ felt that his ideas on observation were not taken seriously by many animal psychologists and that he had a difficult task of making clear that qualitative observations are a necessary step of research. Indeed, in his view the biological community as whole focused alone on measurement and statistical evaluation of data as the ideal of rigorous science.

Originally LORENZ discussed the importance of Gestalt perception without suggesting that his methodological approach might be in need of defense (1948, 1951). In the 50s and 60s, however, he criticized especially behaviorist and American animal psychology in the context of defending his views. LORENZ states, for instance, that the paper *Gestalt perception*... is targeted at American animal psychologists (1958, p256). However, it is not just the case that LORENZ had to defend his methodological views against the influential tradition of behaviorism and other theoretical approaches that had a radically different approach. Some of the remarks in the writings of the 60s and 70s were directed at biolo-

gists in general, because LORENZ felt that his observational approach was considered by many biologists (including German-speaking biologists) as not being really rigorous science. For instance, LORENZ complained about recommendations given by reviewers about a grant application to the Deutsche Forschungsgemeinschaft (German Research Council).

"The grant was approved, but not without the advisory board adding a little benign admonition: care should be taken, lest the investigation lapse into merely being descriptive, "daß die Untersuchung nicht ins Deskriptive *abgleitet.*" (1973, p5).

In addition, LORENZ reacted furiously to the attempt to remove Otto KOENIG as the director of the Austrian Institut für Vergleichende Verhaltensforschung (Institute for Comparative Ethology). Some of the board members stated that the research at the institute was in bad shape. The justification is revealing. On their account, comparative ethology is a amateur program instead of rigorous biology—just making photographs does not presuppose a university degree and the crucial feature of making measurements is not seriously pursued.

"In einer Sitzung des Kuratoriums für das Institut für Vergleichende Verhaltensforschung haben MA-RINELLI und der Physiologe Prof. KMENT den Versuch unternommen, Otto KOENIG hinauszuschmeißen... MARINELLI: Tierhaltung sei nur ein Hobby, die Herstellung von Filmen verlange kein Hochschulniveau (das Wort Hochschulniveau kommt in dem sechs Seiten langen Protokoll der Sitzung 7× vor.) Die vergleichende Verhaltensforschung sei ein Dilletantenprogramm. KMENT (S. 2 unten) gibt eine völlig falsche Vorstellung von unserem Institut und seiner Arbeitsweise. KMENT meint, "entscheidend sei die Messung, man müsse messend vorgehen" und der alte Blödsinn usw." (LORENZ, letter to STRESE-MANN, February 17, 1971; the letter to STRESEMANN includes a copy of the minutes of the board meeting, showing that these accusations were actually raised).

Thus the main target of LORENZ's justificatory efforts is his observational approach, and he felt the need to defend his methodological views against general trends and conceptions in biology and behavioral science. Despite the fact that LORENZ recognized the importance of experiments and quantitative techniques he viewed the scientific contempt for qualitative observation as part of a generally in-

Author's address

Ingo Brigandt, Department of History and Philosophy of Science, 1017 Cathedral of Learning, University of Pittsburgh, Pittsburgh, PA 15260. Email: inb1@pitt.edu creasing disregard for nature. While we may wonder whether Gestalt perception is really as important for ethological observations as LORENZ maintained, we are still not quite clear about the epistemic significance and relation of observation and experiment (ALLEN, forthcoming). LORENZ's achievement is to make clear that we need to have an account of the cognitive mechanisms that are at work in observation to generate scientific knowledge.

Notes

- **1** See BURKHARDT (1999) for a historical comparison of the ways animals and behavior was studied (in museums, in zoos, and in the field) and how this related to and influenced research approaches. BURKHARDT (2003) compares the research practices of LORENZ and TINBERGEN.
- **2** The correspondence between Konrad LORENZ and Erwin STRESEMANN is archived at the Staatsbibliothek zu Berlin, Preußischer Kulturbesitz (Germany) as 'Nachlaß 150 (E. STRESEMANN), Kasten 40'. Any letter from LORENZ that is part of this source is referred to in the paper as 'LORENZ, letter to STRESEMANN'. The letter quoted from above is not dated (it is probably from the end of 1933), 'Blatt 94–97' refers the sheet numbers given by the archive.
- **3** Some of LORENZ's views on epistemology and philosophy of science—in particular as formulated in the *Russian Manuscript*—might appear naive, given the fact that there were much more sophisticated approaches to these topics. At the same time LORENZ developed his philosophical views in the pre-war period in Vienna, philosophers and scientists that were members of the Vienna circle or associated with it developed detailed accounts as to how understand scientific rationality and the justification of empirical knowledge. LORENZ did not pay attention to these developments probably because he and the members of the Vienna circle belonged to different philosophical, scientific, and social-political communities. I will later on

References

- Allen, C. (forthcoming) Is anyone a cognitive ethologists? Biology and Philosophy.
- Ash, M. G. (1998) Gestalt Psychology in German Culture 1890–1967: Holism and the Quest for Objectivity. Cambridge University Press: Cambridge.
- **Brunswik, E. (1934)** Wahrnehmung und Gegenstandswelt: Grundlegung einer Psychologie vom Gegenstande her. Franz Deuticke: Leipzig.
- Brunswik, E. (1937) Psychology as a science of objective relations. Philosophy of Science 4: 227–260.
- Brunswik, E. (1952) The Conceptual Framework of Psychology. Chicago University Press: Chicago.
- **Brunswik, E. (1955)** 'Ratiomorphic' models of perception and thinking. Acta Psychologica 11: 108–109.
- Burkhardt, R. W. (1981) On the emergence of ethology as a scientific discipline. Conspectus of History 1: 62–81.
- Burkhardt, R. W. (1999) Ethology, natural history, the life sciences, and the problem of place. Journal of the History of Biology 32: 489–508.
- Burkhardt, R. W. (2003) Konrad Lorenz, Niko Tinbergen, and

Acknowledgments

I am indebted to Paul GRIFFITHS, Werner CALLEBAUT, Wolfgang SCHLEIDT and Manfred WIMMER for helpful discussions on my work.

point to some useful features of LORENZ's epistemological framework.

- **4** Both were of the same age and working on their Habilitation theses at the University of Vienna. While BRUNSWIK was Karl BÜHLER's assistant, LORENZ took his psychology classes with BÜHLER. BRUNSWIK gave LORENZ comments on the Companion (LORENZ 1935) as regards issues dealing with the psychology of perception (LORENZ, letter to STRE-SEMANN, November 21, 1934). See HOFER (2001) for a discussion of the personal and intellectual relationship between BÜHLER, BRUNSWIK, and LORENZ.
- **5** It is not quite clear what LORENZ means by saying that the Gestalt is just subjective. Gestalt psychologists often assumed that Gestalt perception aims at objective features of the world (see ASH 1998) and LORENZ does not seem to disagree with this in general.
- **6** In a letter from 1936 LORENZ complains that so many physiological chemists are incapable of viewing animals as organic entireties ("Ich kenne aber so viele Physiologische Chemiker, die alle ganz unfähig sind, im Tier eine organische Ganzheit zu sehen", LORENZ, letter to STRESEMANN, October 4, 1936).
- 7 The use of three kinds of justificatory arrows in the diagram is a rough classification. A more detailed and qualified account can be found in the above discussion. Note for instance that an arrow justifying the observational approach refers to some subset of the features peculiar to the observational approach.

the founding of ethology as a scientific discipline. In: Legakis A./Sfenthourakis S./Polymeni R./Thessalou-Legaki M. (eds) The New Panorama of Animal Evolution: Proceedings of the 18th International Congress of Zoology. Pensoft: Sofia, pp. 329–336.

- Burkhardt, R. W. (forthcoming) Patterns of Behavior: Konrad Lorenz, Niko Tinbergen, and the Founding of Ethology. Chicago University Press: Chicago.
- Dewsbury, D. A. (1988) A brief history of the study of animal behavior in North America. In: Bateson P. P. G./Klopfer P. H. (eds) Perspectives in Ethology. Volume 8: Whither Ethology? Plenum Press: New York, pp. 85–122.
- Heinroth, O./Heinroth, M. (1924–1933) Die Vögel Mitteleuropas, in allen Lebens- und Entwicklungsstufen photographisch aufgenommen und in ihrem Seelenleben bei der Aufzucht vom Ei ab beobachtet. 4 volumes. Hugo Bermüller: Berlin.
- Helmholtz, H. (1925) Treatise on Physiological Optics. The Optical Society of America. Ithaca: New York.
- Hofer, V. (2001) Konrad Lorenz als Schüler von Karl Bühler: Diskussion der neu entdeckten Quellen zu den persönlichen und inhaltlichen Positionen zwischen Karl, Bühler,

Evolution and Cognition | **169** | 2003, Vol. 9, No. 2

Konrad Lorenz und Egon Brunswick. Zeitgeschichte 28: 135–159.

- Lillie, F. R. (1911) Charles Otis Whitman. Journal of Morphology 22: xv–lxxvii.
- Lorenz, K. (1931) Contributions to the study of the ethology of social Corvidae. Reprinted in: Lorenz, K. (1970) Studies in Animal and Human Behaviour, Volume I. Harvard University Press: Cambridge MA, pp. 1–56.
- Lorenz, K. (1932) A consideration of the methods of identification of species-specific instinctive behaviour patterns. Reprinted in: Lorenz, K. (1970) Studies in Animal and Human Behaviour, Volume I. Harvard University Press, Cambridge MA, pp. 57–100.
- Lorenz, K. (1935) Companions as factors in the bird's environment: The conspecific as the eliciting factor for social behaviour patterns. Reprinted in: Lorenz, K. (1970) Studies in Animal and Human Behaviour, Volume I. Harvard University Press, Cambridge MA, pp. 101–258.
- Lorenz, K. (1937) The establishment of the instinct concept. Reprinted in: Lorenz, K. (1970) Studies in Animal and Human Behaviour, Volume I. Harvard University Press, Cambridge MA, pp. 259–315.
- Lorenz, K. (1941) Kants Lehre vom Apriorischen im Lichte gegenwärtiger Biologie. Blätter für Deutsche Philosophie 15: 94–125.
- Lorenz, K. (1942) Inductive and teleological psychology. Reprinted in: Lorenz, K. (1970) Studies in Animal and Human Behaviour, Volume I. Harvard University Press, Cambridge MA, pp. 351–370.
- Lorenz, K. (1943) Die angeborenen Formen möglicher Erfahrung. Zeitschrift für Tierpsychologie 5: 235–409.
- **Lorenz, K. (1948)** The Natural Science of the Human Species: An Introduction to Comparative Research. The "Russian manuscript" (1944–1948). Edited from the author's posthumous works by A. von Cranach in 1996. MIT Press: Cambridge MA.
- Lorenz, K. (1950) Part and parcel in animal and human societies. A methodological discussion. Reprinted in: Lorenz, K. (1971) Studies in Animal and Human Behaviour, Volume II. Harvard University Press, Cambridge MA, pp. 115–195.
- **Lorenz, K. (1951)** The role of Gestalt perception in animal and human behaviour. In: Whyte, L. L. (ed) Aspects of Form: A Symposium on Form in Nature and Art. Lund Humphries: London, pp. 157–178.
- Lorenz, K. (1954) Psychology and phylogeny. Reprinted in: Lorenz, K. (1971) Studies in Animal and Human Behaviour, Volume II. Harvard University Press, Cambridge MA, pp. 196–245.
- Lorenz, K. (1958) Methods of approach to the problem of behavior. Reprinted in: Lorenz, K. (1971) Studies in Animal and Human Behaviour, Volume II. Harvard University Press, Cambridge MA, pp. 246–280.
- Lorenz, K. (1959) Gestalt perception as a source of scientific knowledge. Reprinted in: Lorenz, K. (1971) Studies in Animal and Human Behaviour, Volume II. Harvard University

Press, Cambridge MA, pp. 281–322.

- Lorenz, K. (1963a) A scientist's credo. Reprinted in: Lorenz, K. (1971) Studies in Animal and Human Behaviour, Volume II. Harvard University Press, Cambridge MA, pp. 1–13.
- Lorenz, K. (1963b) Do animals undergo subjective experience. Reprinted in: Lorenz, K. (1971) Studies in Animal and Human Behaviour, Volume II. Harvard University Press, Cambridge MA, pp. 323–337.
- Lorenz, K. (1965) Evolution and Modification of Behavior. University of Chicago Press: Chicago.
- **Lorenz, K. (1973)** The fashionable fallacy of dispensing with description. Die Naturwissenschaften 60: 1–9.
- Lorenz, K. (1977) Behind the Mirror. Methuen: London.
- Lorenz, K. (1981) The Foundations of Ethology. Springer Verlag: New York.
- Lorenz, K./Tinbergen, N. (1938) Taxis and instinctive behavior pattern in egg-rolling by the Greylag goose. Reprinted in: Lorenz, K. (1970) Studies in Animal and Human Behaviour, Volume I. Harvard University Press, Cambridge MA, pp. 316–350.
- Nissen, H. W. (1931) A Field Study of the Chimpanzee: Observations of Chimpanzee Behavior and Environment in Western French Guinea. John Hopkins Press: Baltimore.
- **Röell, D. R. (2000)** The World of Instinct: Niko Tinbergen and the Rise of Ethology in the Netherlands (1920–1950). Van Gorcum: Assen.
- Schleidt, M./Schleidt, W. (1958) Kurven gleicher Lautstärke beim Truthhahn (Meleagris gallopavo). Die Naturwissenschaften 45: 119–120.
- **Tinbergen, N. (1932)** Über die Orientierung des Bienenwolfes (Philanthis triangulum Fabr.). Zeitschrift für vergleichende Physiologie 16: 305–334
- **Tinbergen, N. (1935)** Über die Orientierung des Bienenwolfes (Philanthis triangulum Fabr.) II: Die Bienenjagd. Zeitschrift für vergleichende Physiologie 21: 699–716.
- Tinbergen, N. (1951) The Study of Instinct. Oxford University Press: Oxford.
- Tinbergen, N./Kuenen, D. J. (1939) Über die auslösenden und richtungsgebenden Reizsituationen der Sperrbewegung von jungen Drosseln (Turdus m. merula L. und T. e. ericetorum Turton). Zeitschrift für Tierpsychologie 3: 37– 60.
- Watson, J. B. (1908) The behavior of the soddy and nooty terns. Carnegie Publications 103: 187–255.
- Watson, J. B. (1914) Behavior: An Introduction to Comparative Psychology. Henry Holt: New York.
- Whitman, C. O. (1899a) Animal Behavior. Reprinted from the Biological Lectures of the Marine Biological Laboratory, Woods Hole, 1898. Ginn: Boston.
- Whitman, C. O. (1899b) Myths in animal psychology. The Monist 9: 524–537.
- Windelband, W. (1894) Geschichte und Naturwissenschaft. Heitz und Mündel: Straßburg.
- Yerkes, R. M. (1928) The Mind of a Gorilla, Parts 1 and 2. Clark University: Worcester.

Vertical and Horizontal Learning

Some Characteristics of Implicit and Explicit Learning

Introduction

Ever since Michael POLA-NYI introduced the idea of tacit knowing in 1958 much effort has been put into analysing and characterising 'implicit' versus 'explicit' learning, knowledge and memory¹. Implicit knowledge is also known as tacit, automated, non-conscious, procedural and incidental. This is distinguished from symbolically represented knowledge forms (also known as explicit knowledge), which are frequently characterized as able to be expressed in

words (for reviews and discussion, see Berry/Dienes 1993; Reber 1993; Stadler/Frensch 1998).

The first classical study to operationalize the theoretical distinction, known as 'the artificial grammar' paradigm, was conducted by the neuropsychologist REBER (1967). In this original study subjects were presented with short sequences of symbols or letters (for instance pvpxvps and tssxxvps), which followed artificial grammatical rules, known only to the experimenter. Without being informed about the underlying grammar, subjects were asked to remember the sequences. The control group was exposed to sequences of random order. Afterwards, all subjects were told about the existence of an underlying grammar being present in the previously presented letter strings and were asked to classify new sequences in grammatical or nongrammatical strings. Though completely ignorant of the underlying rules subjects formerly exposed to grammatical sequences were significantly

Abstract

Several studies have addressed the differences between implicit and explicit learning. Focus has been on tacit, automated, non-conscious, procedural and incidental learning as opposed to symbolically represented learning, expressed in words. Separation of implicit and explicit learning is generally based on the absence or presence of conscious awareness of the learned material. In this paper, the representational modes of different learning strategies are explored from a phylogenetic perspective, with emphasis on detachment from context. The notion of vertical and horizontal learning is introduced to clarify their applicability to different contexts.

Key words

Implicit learning, explicit learning, conscious awareness, context, abstraction. better at classifying than the control group. However, this superiority was not accompanied by abilities to articulate the rules or strategies by which they separated the strings.

Obviously complex learning might take place without subjects being consciously aware. The learning is simply stored as a difference in response to certain stimuli and not as a conscious rule or strategy subjects control or are capable of volunteering verbally on request.

Though anatomical

(e.g., SQUIRE 1986; SQUIRE/ZOLA-MORGAN 1988, 1991) as well as neuropsychological research (e.g., BERRY 1996; BERRY/DIENES 1993; REBER 1993; STA-DLER/FRENSCH 1998; UNDERWOOD 1996; WEISKRANTZ 1997) unambiguously supports the existence of at least two independent learning systems (e.g., DIGIULIO et al. 1994; SQUIRE/FRAMBACH 1990), disagreement on how theoretically to demarcate implicit from explicit learning has prevailed (e.g., BERRY/DIENES 1993; FRENSCH 1998; SLOMAN 1996; SQUIRE/KANDEL 1999).

Distinctions in terms of consciousness have dominated explanatory attempts, i.e., the presence or absence of awareness, as manifested by the ability or inability to report with confidence on learned material, is crucial to most research paradigms (BERRY 1996; BERRY/DIENES 1993; REBER 1993; STADLER/FREN-SCH 1998; UNDERWOOD 1996; WEISKRANTZ 1997).

As commented by many observers creating a distinction by subjective report might not be feasible,

Evolution and Cognition | **171** | 2003, Vol. 9, No. 2

in that the distinction is operationally unclear (e.g., BERRY 1996; FRENSCH 1998). It is clearly impossible to show that lack of report is due to non conscious awareness of learning. For one, silence with respect to the learned material could come about as a result of the subject's inclination to withhold information he/she is not confident of (for discussions on methodological artifacts in implicit tests see BUCHNER/ WIPPICH 2000)². Moreover self-report is by no means an exhaustive criterion of consciousness (SCHILHAB 2001). (On discussions of self-report and consciousness, see also DAVIS 1997; GOLDMAN 2000; VELMANS 1996).

Irrespective of operational considerations, perhaps it is time to get past the conundrum of consciousness and ask in what way consciousness predicts the function of the different learning systems. Though one cannot argue for one process being superior to another (RAAB 2003), the question of what is achieved by using one strategy rather than the other seems theoretically important (see for instance TAYLOR 2001, WILSON/SCHOOLER 1991). Specifically, what assets and drawbacks characterise implicit versus explicit learning? In what respect are they unique and how they might possibly supplement each other?

To REBER, implicit learning is phylogenetically more original than explicit learning in the sense, that the former antedates the latter and is in general dedicated to ensuring survival. Among other things he predicted distinctive characteristics of implicit learning such as robustness of the memory as well as age and IQ independence.

Here, following the approach recommended by REBER, I characterise implicit and explicit learning within a phylogenetic framework and go one step further to analyse the epistemological qualities (modes of knowledge) specific to each. To characterise implicit and explicit learning epistemologically is to explore which functions they were developed to fulfil. My primary focus will be on contextual matters in learning. The paper focuses on the issue of detachment from reality by using either mode of learning. Specifically, how explicit learning condenses reality will be considered. Based on this discussion I will introduce the notion of vertical and horizontal learning to capture the specific attributes of implicit and explicit learning respectively. Vertical learning (i.e., depth) refers to implicit learning understood as contextual learning based on relational aspects, such as the connectivity of available information in learning situations. Horizontal learning (i.e., width) refers to explicit learning as context free, abstract learning. The concept of vertical and horizontal learning clarifies what separates implicit and explicit learning as regards evolutionary adaptation. To qualify claims on the consciousness-based distinction of implicit and explicit learning REBER (1992, 1993) recently argued for an evolutionary stance on the different qualities of the learning systems.

To Be Explicit is to Be Context-Free

Explicit learning is stored as knowledge that can be represented deliberately by words or other symbols. The symbolic character of explicit learning makes it possible to 'translate' knowledge obtained in one 'language' into different vocabularies by being context free (DEACON 1998).

To exemplify; we can talk of the capital of France as Paris, as a red mark on a map, as a fixed letter string, as the location which accommodates the Eiffel Tower and the Triumphal Arc or as the city in which the final stage of the Tour de France ends (among other designations). All of these designations have Paris as their referent. Though Paris is determined differently from one conception to the other, by disregarding the information specific to any particular interpretation, one finds the referent 'Paris' to be intrinsic to all.³

Hence 'Paris' is implied in all of the above-mentioned descriptions, and can replace any of them without changing the meaning of the sentence. This transparency entails that we are equally capable of treating the notion in the language of geography as in the language of professional bike racing. Thus, the symbolic nature of explicit knowledge shows extensive manipulability by being detached from context.

Two Senses of 'Detachment from Context'

What is meant by detachment from context? Taken at face value, the assertion is false. Paris is always defined as part of a greater whole, (the bicycle race, capital of France, scene of the French Revolution etc.) though the symbolic representation appears to grasp the essence of Paris. When dealing with different understandings of Paris we choose perspective. Initially, as we came to know of Paris some elements (variables) or contingencies seemed relevant to our particular idea of Paris, while others were concurrently discounted as irrelevant. For instance if interested in bicycling, watching the Tour de France

Evolution and Cognition | **172** | 2003, Vol. 9, No. 2
might imply focus on the pavements of Champs Élyseés and general weather conditions of Paris, while historical events and memorials are ignored.

What is true, however, of explicit learning is detachment from context understood as either condensation of reality to, (a) variables of one's choice or (b) simplification of parameters in learning situation by idealisation⁴.

Why is that? Firstly, processing of symbolically represented knowledge (that is knowledge that can be mentally manipulated) is restricted. The number of elements that can possibly be stored in shortterm memory (working memory) seems limited. For instance, it is possible to retain only 3-4 colours or orientations in visual working memory at a time (VOGEL/WOODMAN/LUCK 2001). Limitations also exert themselves in the number of digits in telephone numbers one can hold in memory (SQUIRE/KANDEL 1999). In practice the actual number of variables relevant to any given learning situation is constrained by this limited capacity. Thus keeping in mind certain features at the expense of others is inevitable (see also DREYFUS/DREYFUS 1986, 1998). Accentuating certain parameters takes place at the expense of an infinite number of potential relevant parameters. Crudely put the production of knowledge simultaneously results in the production of non-knowledge (HOFFMEYER 1984).

Secondly, the very interpretive process of abstract thought fragments real life entities into nonexistent pieces of thought, which are mentally manageable but lack one-to-one correspondence to actual referents. To exemplify; when speaking of 'Granny Smith' as a type of green apple, one determines all Granny Smith apples as green. Naturally, this seems trivial. But if one scrutinises apple skins, the use of the notion 'green' seems unjustified. The actual colour of Grannies might change from apple to apple, might be a mixture of a large variety of colours that exist exclusively in conjunction with the angle of observation, the light intensity and the viewing distance. Postulates about the colour of apple skins ease the burden of communication, but 'green Grannies' as a notion, i.e., abstracted from actual apples understood as a single colour, easily disclosed by scrutiny, do not exist in reality. By denoting, for example, the colour green, verbal (or symbolic) identification demarcates referents in the world that do not actually exist detached from context. If Martians, that neither defecated nor ate were to visit Earth (as described by DAWKINS 1998) and see apples for the first time they would probably not respond to the same features as we do. Perhaps they would even question whether apples were separate from branches or leaves, for the simple reason that they were indifferent to the consequences of such information, i.e., being non-eaters.

Some might hold that the example about the colour of apple skins is misleading, since, this argument might hold with respect to 'green' but not with respect to natural kinds such as gold and tigers. However, one can rebut the objection by the claim that definitions of gold and tigers still imply an accentuation of certain criteria at the expense of others. By convention we apply criteria and thereby denote gold or tiger (KURODA 1998). By stressing certain features (for instance greenness) of apples the notion of apple is detached from the context in which any apple actually appears. As such, symbol use virtually constructs reality in a radical way.

While condensation of reality in the sense of reduction of numbers of variables might principally be overcome by improved cognitive abilities such as increased memory span (COLLINS 2001), detachment of context by simplification seems insurmountable. To fully comprehend the implications of creating cognitive schemas for objects think of the familiar table. Normally, we encounter no difficulty in defining where the table begins or ends although molecularly speaking (at the micro level) 'ending' and 'beginning' is meaningless. Incessant exchanges of molecules blur the boundaries between table and space. After all, it is meaningful to presuppose the existence of ending and beginning of tables simply because human perception operates at the macro level (at which Newton's laws take effect) (LAKOFF 1990).

Implicit Learning and Condensation of Reality

So far I have claimed that explicit learning is subordinate to (or is characterised by) detachment of context by 1) reducing the number of variables in a given learning situation due to perspective taking and 2) accentuating certain features of the variables and ignoring others. How does the question of detachment from context relate to implicit learning?

Of interest is whether implicit learning, by being responsive to relations between many parameters without involving awareness, avoids contextual detachment. My answer is no. But implicit learning is significantly less detached from reality by there simply being less interpretation, a crucial distinction when separating explicit from implicit learning.

Evolution and Cognition | 173 | 2003, Vol. 9, No. 2

What supports this contention? First, explicit learning involves a significant reduction of parameters in mental representation, which I refer to as "condensation of reality". But research paradigms demonstrating implicit learning are typically complex in the sense of concealing a huge number of parameters essential to solving the task in a large web of information, which makes explicit strategies impossible to employ. For instance VINTER/PERRUCHET (2000) reported on implicit drawing behaviour in children aged 4-10 years. To confront criticism related to explicit knowledge criteria, the researchers developed a paradigm named a "neutral parameter procedure" and deliberately diverted participant's attention to test clues to avoid the use of explicit knowledge obtained during the test procedure. Experimenters specifically attempted to blur conscious awareness of connections between any two variables. Still, by responding appropriately, the subjects seemed capable of catching contingencies between many variables, while being perfectly unaware of doing so.

The same applies to the artificial grammar test. When subjects were asked to explain how they discriminated between grammatical and nongrammatical sentences, they either denied consciously following any rule or they gave reasons irrelevant to solving the task.

Similar results are obtained with computer simulations of sugar fabrication (BERRY/DIENES 1993). Here people were asked to attain a specific amount of sugar (in tons) by varying different parameters interconnected by a pattern created by the experimenters, but opaque to the subjects. After a number of trials they succeeded. When subjects were subsequently asked to verbalise which rules they applied to obtain the outcome and these rules were followed by novices the end result was strikingly poor. Obviously, the articulated rules do not mirror the rules actually applied. Somehow knowledge residual to the volunteered information remains to be revealed. In support of this are studies within the sugar production paradigm that focus on how to make explicit what could originally be learned only implicitly. By pinpointing the connections between dependent variables or reducing the number of parameters responsible for a successful outcome, subjects suddenly gained insight they could subsequently verbalise. Evidently, to switch from implicit to explicit modes of learning involves both accentuating elements that determine the task and reduction of the number of elements (DIENES 1993).

Implicit Learning and the Evolutionary Stance

The premise that implicit learning is more congruent with reality than explicit learning as regards representation can be explained theoretically. From an evolutionary perspective, for social organism like humans the physical as well at the social environment is enmeshed in information of potential importance to survival (BYRNE 1995). The mere fact of balancing the surface can be realised only by taking in a large amount of information (COTTERILL 2002). Likewise, navigation in the social world, originally adapted to interpret tribal relations and coalitions in prehistoric society (WRANGHAM et al. 1994), now entails coping with a multitude of modern social conventions and regulations (COLLINS 2001) such as, for example, table manners and traffic rules. Information important to survival does not come about in isolated incidents, but relates to cues in the environment.

Environmental complexity clearly explains why implicit learning is less detached from context with respect to condensating reality. But the question remains, does this hold for (a) detachments caused by simplification and (b) does not any interpretation involve detachments irrespective of whether it results in simplifcations or distortions of the original source of information? Of course the answer is affirmative. Over time, natural selection favoured perception and learning of certain stimuli and environmental contingencies, which naturally entails perspective-taking followed by simplification or distortion of information. Though the difference in contextual detachment between implicit and explicit learning is a matter of degree, the consequences are significant. Explicit learning, however, emerged as an expansion to an already well functioning system (see also EN-NEN 2002). The contribution of the new system was primarily to control learning capabilities, including the deliberate selection of information to be learned and the intentional application of the obtained (GIBSON/INGOLD knowledge 1993). Though the explicit learning system evolved to profit from intentionel manipulation of information, it emerged to increase survival.

To understand in what respect implicit learning is closer to context we must return to evolutionary considerations. Implicit learning registers a greater number of variables, as well as their intricate pattern, 'invisible' to the explicit capacity, because it seems appropriate to survival. But why? This appears due to that implicit learning was developed

Evolution and Cognition | 174 | 2003, Vol. 9, No. 2

as a system meant to increase the probability of immediate survival. Organisms responsive to information that would be potentially harmful or stimulating would be favoured over organisms with less well adapted monitoring capabilitites. This simple fact is crucial to the evolutionary invention of the learning system (CHURCHLAND 1986; DENNETT 1996). For the implicit learning device to acknowledge and categorize encounters into harmful and productive recurrence of events are conditional. Monitoring seemingly neutral stimuli signaling the occurrence of either harmful or productive stimuli (central to almost all sorts of conditioning processes) increases the survival capacity. If an organism can anticipate events by monitoring cue signals, it is one step ahead of reality. Therefore context matters. In comparison, a learning device worked out to register information which is seldom repeated or barely had any consequences would be outcompeted and therefore not implemented in the gene pool of the generations to come. Learning systems focused on environmental regularities will tend to be sensitive to a broad range of stimuli, combining and storing different stimulus modalities to obtain as much information as possible. Incidently, this explains why implicit learning is applicable to all our senses. Being receptive to any event however insignificant, because of its predicitive powers is to focus on relations and interconnectivity of variables.

The crux of the matter is as follows: Implicit learning capabilities emerged to monitor intricate patterns of information of survival value to the organism. To register relations and interdependency of stimuli is to learn about context. Hence, implicitly learned material retain context and becomes correspondently committed to specific learning situations to a greater extent.

Explicit and implicit learning systems interpret the environment differently as regards detachment from context. Implicit learning relates to information as pieces in a jigsaw puzzle relate to the picture they represent. By using an implicit learning strategy one might not identify all relationships due to a number of pieces being missing, but one might still

get a fairly good representation of the overall picture. On the contrary, using an explicit learning strategy corresponds to choosing a few significant pieces to get a general idea (by educated guesses) of the whole by disregarding a large amount of less significant pieces of information. In the explicit condition if one chooses pieces of sky, ground and corners the impression applies to a multitude of pictures. Whereas in the implicit condition the extra pieces of for example birds, houses and telephone poles fill in relations to determine this particular picture.

Horizontal and Vertical Learning

To summarise; implicit learning is adapted to monitor intricate patterns whereas the benefit of explicit learning is to abstract material from relations and contexts, to gather and store particular information in a universal format that can be applied in a new context. The independence of actual events makes the learned material mentally manageable which can eventually be passed on to other individuals.

For any organism, the first priority is to stay alive. Hence abilities to divide immense amounts of information into essential (life preserving) and non essential (unimportant to life preservation) information will be aimed at stimuli that show some sort of repetition and form part of a recognizable pattern, specific to a learning incident. Disclosing structures that are vertically oriented becomes essential and results in the learning not being applicable to different contexts.

On the other hand, the explicit learning system was developed on top of the implicit system, not as an alternative solution to replace implicit learning devices. Thus, explicit learning, that supplements an already well functioning implicit learning system, is adapted to obtain information irrespective of contextual relations. The explicit learning system can be applied to a wider range of information processing. For example, it has no need for stimuli to be constant or repetitive, in order to operate. In many instances, one exposure to a particular stimulus is sufficient for later storage. Because of being widely independent of context it can also be applied to different contextual settings which makes it horizontally oriented. Explicit learning capabilities (especially context independency) were beyond any doubt fundamental to the cultural progress of humans as viewed by ad-

Author's address

Theresa S. S. Schilhab, Learning Lab Denmark, Danish University of Education, Emdrupvej 101, DK-2400 Copenhagen NV. Email: theresa@lld.dk vanced tool making (DAVID-SON/NOBLE 1993; NOBLE/ DAVIDSON 1996). The superiority of human intelligence, in particular its versatility, is due to the sophistication of the explicit learning device (e.g.; BENNETT 1988, 1997).



The spatial metaphor of horizontal and vertical learning implies a substantial difference as regards the applicability of implicit versus explicit knowledge. Knowledge resulting from implicit learning is highly applicable to similar learning situations (because it effectively manages complexity) but applicability decreases with increasing dissimilarity (implying that context is essential).

In comparison, knowledge resulting from explicit learning is applicable to a much wider range of learning situations, due to the symbolic nature of its representation. As explicit knowledge stresses the specific (and few) aspects of learning situations it is appropriate for a large number of contexts. I claim that the capacity to transfer from one context to another is largely gained by leaving out of account relations in which the learning takes place.

Figure 1 is meant to illuminate these characteristics. However, I think vertical (implicit) learning always operates along with horizontal (explicit) learning, thus, I do not claim that the efficiency of vertical learning actually decreases to zero, rather that at some point horizontal learning becomes more efficient. In this respect the two ways of learning complement each other, as this applies to humans.

Concluding Remarks

The assets and drawbacks of implicit and explicit learning are telling. Obviously evolution opted for two different epistemological strategies: The first strategy is to secure immediate survival by vertical learning that is receptive to stimuli in their contextual settings thereby making it possible for humans to navigate safely in specific environments. Secondly, on top of the well-established implicit learning horizontal learning emerged by which abstract thinking and the ability to transfer experiences from one learning context to another, and eventually from one individual to another, became possible.

Notes

- 1 Learning is the acquisition of knowledge, and memory the retention of that knowledge. As such learning, knowledge and memory are different phenomena, but can hardly be totally separated in practice. However to the present discussion I will refer to all three aspects unless clearly stated.
- **2** To circumvent this difficulty researchers often make use of questionnaires to ask specific questions thus minimising any doubt subjects could have. However, accentuation of parameters and connections crucial to the task raises the risk of priming subjects. Moreover, asking subjects specific

References

- Bennett, J. (1988) Thoughtful brutes. American Philosophical Proceedings and Addresses Association 62(4): 197–210.
- Bennett, M. R. (1997) The Idea of Consciousness. Harwood Academic Publishers: Amsterdam.
- Berry, D. C. (1996) How implicit is implicit learning? In: Underwood, G. (ed) Implicit Cognition. Oxford University Press: Oxford, pp. 203–225.
- Berry, D. C./Dienes, Z. (1993) Implicit Learning. Lawrence

questions means articulating the task into details that may not have played any part in solving the task originally. Thus subjects that cannot confirm to solve the task as suggested by the questionnaire are falsely categorised as unaware of the learned material. However, the possibility exists that subjects were fully aware, yet of parameters different from those presented in the questionnaire (DIENES 1993).

- **3** Please notice that nothing ontological (real existence) about Paris is inferred.
- 4 As emphasized by one reviewer, this sort of reduction is not to be confused with philosophical "Reductionism" as a heuristic for finding satisfactory explanations for phenomena.

Erlbaum Associates Ltd: East Sussex.

- **Buchner, A./Wippich, W. (2000)** On the reliability of implicit and explicit memory measures. Cognitive Psychology 40: 227–259.
- Byrne, R. W. (1995) The Thinking Ape. Oxford University Press: Oxford.
- Churchland, P. S. (1986) Neurophilosophy. MIT Press: Cambridge.
- Collins, H. M. (2001) What is tacit knowledge? In: Schatzki, T. R./Cetina, K. K./Savigny, E. v. (eds) The practice Turn in

Evolution and Cognition | 176 | 2003, Vol. 9, No. 2

Contemporary Theory. Routlegde: London, pp. 107–119. Cotterill, R. (2002) Bevidsthed, intelligens og kreativitet. Ak-

tuel Naturvidenskab 3: 40–43. Davidson, I./Noble, W. (1993) Tools and language in human evolution. In: Gibson, K. /Ingold, T. (eds) Tools, Language and Cognition in Human Evolution. Cambridge University Press: Cambridge, pp. 363–388.

Davis, H. (1997) Animal cognition versus animal thinking: The anthropomorphic error. In: Mitchell, R. W./Thompson, N. S./Miles, H. L. (eds) Anthropomorphism, Anecdotes and Animals. State University of New York Press: Albany, pp. 335–347.

- Dawkins, M. S. (1998) Through Our Eyes Only? Oxford University Press: Oxford.
- **Deacon, T. (1998)** The Symbolic Species: The Co-Evolution of Language and the Human Brain. Norton: New York.
- Dennett, D. C. (1996) Kinds of Minds. Weidenfeld & Nicholson: London.

Dienes, Z. (1993) Implicit concept formation. In: Berry, D. C. /Dienes, Z. (eds) Implicit Learning. Lawrence Erlbaum Associates Ltd: East Sussex, pp. 37–61.

- **Digiulio, D. V./Seidenberg, M./O'leary, D. S./Raz, N. (1994)** Procedural and declarative memory: A developmental study. Brain and Cognition 25: 79–91.
- Dreyfus, H./Dreyfus, S. (1986) Mind Over Machine. Free Press: New York.

Dreyfus, H./Dreyfus, S. (1998) The challenge of Merleau-Ponty's phenomenology of embodiment for cognitive science. Skriftserie, Institut for filosofi 9: 1–25.

Ennen, E. (2002) The basal ganglia and the cognitive unconscious. Evolution and Cognition 8(2): 123–129.

Frensch, P. A. (1998) One concept, multiple meanings. In: Stadler, M. A. /Frensch, P. A. (eds) Handbook of Implicit Learning. Sage Publications: Thousand Oaks, pp. 47–104.

- Gibson, K./Ingold, T. (eds) (1993) Tools, Language and Cognition in Human Evolution. Cambridge University Press: Cambridge.
- Goldman, A. I. (2000) Can science know when you're conscious? Journal of Consciousness Studies 7(5): 3–22.
- Hoffmeyer, J. (1984) Hovedet i naturen. Rosinante: Viborg.

Kuroda, W. (1998) Other minds. In: Nagasaka, F. G. (ed) Japanese Studies in the Philosophy of Science. Kluwer Academic Publishers: London, pp. 7–19.

Lakoff, G. (1990) Women, Fire and Dangerous Things. University of Chicago Press: Chicago.

- Noble, W./Davidson, I. (1996) Human Evolution, Language and Mind. Cambridge University Press: Cambridge.
- Raab, M. (2003) Decision making in sports: Influence of complexity on implicit and explicit learning. International

Journal of Sport and Exercise Psychology 1: 310–337.

- **Reber, A. S. (1967)** Implicit learning of artificial grammars. Journal of Verbal Learning and Verbal Behavior 6: 317–327.
- **Reber, A. S. (1992)** The cognitive unconscious: An evolutionary perspective. Consciousness and Cognition 1: 93–133.
- Reber, A. S. (1993) Implicit Learning and Tacit Knowledge. Clarendon Press: Oxford.
- Schilhab, T. (2001) Man and the mirror. Evolution and Cognition 7(2): 178–186.
- Sloman, S. A. (1996) The empirical case for two systems of reasoning. Psychological Bulletin 119(1): 3–22.
- Squire, L. R. (1986) Mechanisms of memory. Science 232: 1612–1619.
- Squire, L. R./Frambach, M. (1990) Cognitive skill learning in amnesia. Psychobiology 18(1): 109–117.
- Squire, L. R./Kandel, E. R. (1999) Memory. Scientific American Library: New York.
- Squire, L. R./Zola-Morgan, S. (1988) Memory: Brain systems and behavior. Trends in Neuroscience 11: 170–175.
- Squire, L. R./Zola-Morgan, S. (1991) The medial temporal lobe memory system. Science 253: 1380–1386.
- Stadler, M. A./Frensch, P. A. (eds) (1998) Handbook of Implicit Learning. Sage Publications: Thousand Oaks.
- **Taylor, E. W. (2001)** Transformative learning theory: A neurobiological perspective of the role of emotions and unconscious ways of knowing. International Journal of Education 20(3): 218–236.
- **Underwood, G. (ed) (1996)** Implicit Cognition. Oxford University Press: Oxford.
- Velmans, M. (1996) What and where are conscious experiences? In: Velmans, M. (ed) The Science of Consciousness. Routledge: London, pp. 181–196.
- Vinter, A./Perruchet, P. (2000) Implicit learning in children is not related to age: Evidence from drawing behavior. Child Development 71(5): 1223–1240.
- Vogel, E. K./Woodman, G. F./Luck, S. J. (2001) Storage of features, conjunctions, and objects in visual working memory. Journal of Experimental Psychology: Human Perception and Performance 27(1): 92–114.
- Weiskrantz, L. (1997) Consciousness Lost and Found. Oxford University Press: Oxford.
- Wilson, T. D./Schooler, J. W. (1991) Thinking too much: Introspection can reduce the quality of preferences and decisions. Journal of Personality and Social Psychology 60(2): 181–192.
- Wrangham, R. W./McGrew, W. C./De Waal, F. B. M./Heltne, P. G. (eds) (1994) Chimpanzee Cultures. Harvard University Press: Cambridge.

Of Brains and Minds

A Neurobiological Treatise on the Nature of Intelligence

Brain, Mind and Reality

Organisms are faced during their lives with an immense variety of problems. ranging from physical purely ones, such as changes in climate or geomorphic disturbances, to organismspecific problems related to food supply, predation, homeostasis, reproduction, etc. In order to enhance their chances of survival, organisms have to find adequate solutions for the problems with which they are confronted, for any of them could easily be fatal. Problem solving, in other words, is an essential dy-

Abstract

All organisms are faced during their lives with an immense variety of environmental challenges and organism specific problems, for which they have to find adequate solutions, in order to survive. In the present essay, biological intelligence is considered to be a correlate of the problem solving capacity of an organism, manifesting itself in the amount of neural information that is used in perceiving and interpreting the external (and internal) world. Consequently, biological intelligence can be conceived as to reflect the temporal and spatial complexity of the species niche. In accordance with this view, the hypothesis is put forward that in higher organisms the complexity of the neural (micro)circuitry of the cerebral cortex is the neural correlate of the brain's coherence and predictive power, and, thus, a measure of biological intelligence.

Key words

Biological intelligence, information processing, consciousness, problem solving, cognition, cerebral cortex, evolutionary epistemology.

namic survival mechanism, evolved to cope with disturbances in the ecological equilibrium. It can therefore be looked upon as an adaptive capacity enabling organisms to adjust themselves to one another and to their physical environment.

The kind of problems with which organisms are confronted, however, and their relative significance, varies from one species to another, according to the ecological niche or adaptive zone that it occupies. These specific environmental challenges form the selection pressures that have given rise to the evolution of species-specific neural mechanisms and action patterns. In fact all organisms are problem solvers, and the problem solving capacity of a species reflects the temporal and spatial complexity of its environment. Consequently, the ability to solve problems will manifest itself in all those situations in which subjects are required to respond adequately to novel objects and changing circumstances, as well as in situations in which successful adaptation involves the detection of an appropriate response to regularities in the external world or the formation of rules and hypotheses (POPPER 1982; HODOS/CAMPBELL 1990; MACPHAIL 1993; SHETTLE-WORTH 1998). Environmental adaptation, therefore, can be considered to be the primary function

of problem solving in that it serves as a preeminent mechanism for survival.

The organism's adaptability, however, is but one aspect of fitness. Free-moving organisms, for example, can actively explore their environment, and thus generate new selection forces that can modify the structures involved. MAYR (1982, p612) even argues that 'many if not most acquisitions of new structures in the course of evolution can be ascribed to selection forces by newly acquired behaviors'. This suggests that in highly complex organisms, such as mammals, behavior rather than environmental change may be the major driving force for evolution at the organismal level. However, this

Evolution and Cognition | 178 | 2003, Vol. 9, No. 2

does not detract from the fact that all organisms, whether they are simple reflex automata or active and complex explorers, are above all concerned with keeping track of their local spatio-temporal environment, as part of their struggle for existence. Since sensory information processing and the ability to model reality (or certain parts of it) are essential components in this process, our idea of problem solving seems to correspond reasonably well to the notion of intelligence.

However, the common use of the term 'intelligence' applies not only to processes involving complex information processing (perception) but includes operations of the 'mind' as well (see, e.g., DENNETT 1983; MACPHAIL 1985, 1993). It means that if to a group of organisms no thoughts, intentions, expectations and the like can reasonably be attributed, they are considered to be creatures lacking in intelligence. In order to avoid such subjective criteria, intelligence in the present essay is defined as the problem solving capacity of a species. It implies that intelligence is not a quality restricted to the functional domain of complex neural structures, but can in principle be attributed to all organisms, from amoeba to man.

Although each organism is equipped with an execution potential which enables it to cope with a variety of problems in a specified environment, there are, of course, tremendous differences in the problem solving capacity among living beings, and thus in biological intelligence. Most of these differences are connected with the functional characteristics of the organism's perceiving and executing apparatus. A coelenterate, for instance, with its diffuse nerve fiber network, has an action pattern which is of quite a different order of magnitude as compared to the rich behavioral repertoire of primates, with their highly evolved central nervous system. Therefore, it is only meaningful to compare the intellectual capacities of species when they have certain basic features in common.

The mammalian brain can be considered to be such a structure, in that it is a multimodal integration system composed of a highly efficient hybrid device consisting of analogue neural units to process information and a digital wiring system for information transmission. In these highly organized animals information from the external world passes through three distinct stages or systems: a sensory transducer system, a perceptual input system (or systems) and finally a central cognitive system. During this 'journey' the otherwise overwhelming amount of sensory information is selected, analyzed, integrated and stored in accord with the species' attention and its needs and interests. It means that the picture that an animal has of its external world depends on (i) the quality of its sense organs, (ii) its information processing capacity and (iii) its informational and emotional states of mind. Consequently, 'world pictures' of animals must perforce differ from each other, and can be looked upon as highly individual representations of the external world. Hence, it is appropriate to speak of speciesspecific perceptual worlds (JERISON 1973, 1991) or models of reality (WUKETITS 1986, 1990; RIEDL 1987; PLOTKIN 1994).

Neural Worlds and Real Worlds

The idea of a species-specific model of reality corresponds, to a certain extent, with KANT's assumption that the world as we know it is our interpretation of the observable facts in the light of theories that we ourselves invent (POPPER 1958; see also VOLLMER 1992). It means that models of reality, at least those of higher vertebrates, are related to both the external and the internal worlds and that it assumes the existence of some knowledge in the form of dispositions and expectations (POPPER 1972). The more complete and reliable these knowledge-based specific models are, the better the chance of survival. It enables the animal to make better predictions, especially predictions relating to features or situations which do not occur in stereotyped patterns.

Though states of the brain represent states of the external world we do not perceive reality precisely as it is. The epistemological question now is how to address the relationship between the neural world and the real world. Perhaps representational models in the brain are roughly like a map in the sense that internal, abstract relationships map onto the external relationships between various categories in the world. According to this theory, put forward by CHURCHLAND and CHURCHLAND (2002), brains develop high-dimensional maps, the internal distance relationships of which correspond to the similarity relationships that constitute the categorial structure of the world. The rough and low-dimensional analogy is the road map of a city, in which the real spatial relationships between roads are represented in the relationships between road-lines on the paper map. Just as road maps come in varying degrees of reliablity and detail, so brain models of the external world map the categorial and causal structure of the world with varying degrees of reliablity and detail. A frog's brain maps less of the categorial struc-

Evolution and Cognition | 179 | 2003, Vol. 9, No. 2

ture of the world than a monkey's brain and an infant's brain maps less than an adult's brain. But unlike two-dimensional road maps, the neural maps of reality will be multi-dimensional, probably very high-dimensional maps (CHURCHLAND 1996; GLYMOUR 2001).

The coherence and predictive power that representational models of reality enjoy is explained by biological evolution. The better and faster the brain's predicitve capacities relative to the organism's modus vivendi, the better its chances of survival and reproduction. In the broadest terms, the solution found by evolution to the problem of prediction is to modify execution programs by sensory information. The value of the sensory input is greater if it can signal organism-relevant features and causal regularities between events. To achieve this, the system needs neural cell clusters that are interposed between the sensory system and the motor system to find and embody higher order regularities. According to CHURCHLAND and CHURCHLAND (2003), the richer the interposed neural resources, the more sophisticated the statistical capacities and the greater the isomorphisms achievable between the brain's categorial/causal maps and the world's categorial/causal structures. Higher degrees of isomorphism lead to more reliable models of the world. As we cannot directly compare representational models and the world modelled, predictive success is the measure of fidelity and the guide to the need for model revision. The reality-appearance distinction ultimately rests on comparisons between the predicitve merits of distinct representational models; the better the model's predicitve profile the closer it is to the truth.

Neural Substrates of Intelligence

If we now assume that biological intelligence in higher organisms is the product of processes of complex sensory information processing and mental faculties, responsible for the planning, execution and evaluation of intelligent behavior, variations among species in intelligence must in principle be observable in the neural substrate. Before attempting to determine the underlying neural mechanisms of intelligence, we should have in mind a specific biological entity towards which to direct our attention. Conceiving biological intelligence as the capacity of an organism to construct an adequate model of reality, implies that the spectrum of inquiry may range from the sensory receptor system to behavior in its broadest sense (that is

to say, overt activity as well as internal homeostatic action). Usually, however, valid comparisons at the extremes of the spectrum i.e., at the level of sense organs and complex behavior patterns, respectively, are difficult to make in view of the very great sensorimotor differences that exist among species (see, e.g., PIRLOT 1987; MACPHAIL 1993, MACPHAIL/ BOLHUIS 2001; ROTH/WULLIMANN 2001). How to compare in mammals, for example, the sensory capacity of diurnal monkeys with stereoscopic vision with that of nocturnal echolocating bats, or the learning ability of terrestrial shrews with that of marine dolphins? Differences in intelligence may in fact be uncorrelated with measurable differences in overt behavior, nor are such differences implicit in many learning situations, since both activities depend on the behavioral potential of the organism as well as on its internal state (attention, motivation, etc.; a starving rat, for instance, is probably not more intelligent than a satiated one!).

To avoid these formal problems, one should instead investigate the 'general-purpose' (neo)cortical areas, where both perception and instruction take place. It is the organism's neural substrate where the external world is interpreted and modelled, where concepts are formed and hypotheses tested, in short, where the physical world interacts with the mind. Since the primary function of the brain is to adequately interact with the external world, brain function can be most readily characterized by the manner in which the brain senses the physical environment and how it responds to it by generating motor actions. From experimental and theoretical studies it has become evident that the brain is a distributed parallel processor where most of the sensory information is analyzed in parallel involving large neuronal networks (FREEMAN 1975; BAL-LARD 1986; ZEKI/BARTELS 1999). The principle of parallel processing implies that activities of ordered sets of nerve cells can be considered to be mathematical vectors (or tensors). An important aspect of this vector approach is that it focuses on the explanation of brain function in terms of neuronal networks, and that it is therefore compatible with the modular and hierarchical organization of the brain.

Despite these major developments to explain brain function in terms of tensors, the attempts have so far been confined to sensorimotor operations (PELLIONISZ 1988), whereas no current theory successfully relates higher brain functions to details of the underlying neural structure. This is hardly surprising in view of the enormous functional complexity of the brain, especially that of higher verte-

Evolution and Cognition | **180** | 2003, Vol. 9, No. 2

brates. Instead, one may ask a more general question, one which is related to the evolution, development and function of neural information processing: does the brain operate according to a general mechanism or principle in processing sensory signals and stored information, despite the manifold differences in brain subsystems and their interconnections? After all, brains must be able to process extremely complex information, but they must also have a simple enough underlying organization to have evolved by natural selection (GLASS-MAN 1985).

Design Principles of Neuronal Organization

If the neuron can be regarded as the 'atomic' unit of function in the nervous system, then the 'molecular' unit of information processing is in a way akin to the neuronal network. In particular, the mammalian neocortex has been found to be uniformly organized and to be composed of such neural processing units (or modules) interacting over fairly short distances (SZENTÁGOTHAI 1978; HOFMAN 1985, 1996; CHERNIAK 1995; MOUNTCASTLE 1997, 1998; BUXHO-EVEDEN/CASANOVA 2002a). It appears that the module for information processing in the neocortex consists of a functional neuronal network with a columnar structure that has the capability of quite sophisticated spatial-temporal firing patterns. These processing units operate as prewired neural assemblies where individual neurons are configured to execute fairly complex transactions. Their widespread occurrence, furthermore, qualifies them to be considered as fundamental building blocks in neural evolution (for reviews see HODOS 1982; MOUNTCAS-TLE 1997; BUXHOEVEDEN/CASANOVA 2002b).

We are beginning to understand some of the geometric, biophysical and energy constraints that have governed the evolution of these neuronal networks (e.g., CHKLOVSKII/SCHIKORSKI/STEVENS 2002; KLY-ACHKO/STEVENS 2003; LAUGHLIN/SEJNOWSKI 2003). To operate efficiently within these constraints, nature has optimized the structure and function of these processing units with design principles similar to those used in electronic devices and communication networks. In fact, the basic structural uniformity of the cerebral cortex suggests that there are general architectural principles governing its growth and evolutionary development (CHERNIAK 1995; RAKIC 1995; HOFMAN 1996, 2001a).

It is now well established that the cerebral cortex forms as a smooth sheet populated by neurons that

proliferate at the ventricular surface and migrate outwards along radial glial fibers, forming distinct neuronal networks that are organized in columnar arrays stretched out through the depth of the cortex (Leise 1990; Malach 1994; Krubitzer 1995). It has been postulated that these neural processing units, or modules, have spatial dimensions depending on the number of local circuit neurons and that both the number and size of cortical modules increase with increasing brain size (HOFMAN 1985, 1991; PROTHERO 1997). Scaling models, furthermore, indicate that the differences in modular diameter among mammals is only minor compared to the dramatic variation in overall cortex size. Thus it seems that the main cortical change during evolution has presumably been an increase in the number, rather than the size of these neural circuits.

New light has been thrown on this matter by linking the modular concept to the idea of corticalization, in an attempt to explain the augmented information processing capacity of the mammalian cerebral cortex (HOFMAN 1985, 2001a). According to this theory, the neural processing units or modules of the cortex are wired together so as to form complex processing and distribution units, having spatial dimensions depending on the species' degree of corticalization. As a result, the structural complexity of these processing units increases with the evolutionary expansion of the cerebral cortex, and with that, their functional capacity. Analogue organizational principles are known from computer technology, where the achievements of complex systems are found to depend on emergent mechanisms (cf. species-specific neural processing units) rather than on the quantitative addition of units and interconnections with the same properties as found in simpler systems (CHANGIZI/MCDANNALD/WIDDERS 2002; COCHRANE/ WINTER/HARDWICK 1998).

It has become evident that cortical modules integrate at higher levels of information processing, as a result of the hierarchical organization of the brain, thus enabling the system to combine dissimilar views of the world. It implies that if we seek the neural basis of biological intelligence, including mind-like properties and consciousness, we can hardly localize it in a specific region of the brain, but must suppose it to involve all those regions through whose activity an organism is able to construct an adequate model of its external world, perhaps it may even encompass the entire neo- and subcortical network.

Evolution and Cognition | 181 | 2003, Vol. 9, No. 2

Biological Limits to Information Processing

If the ability of an organism to process information about its environment is the driving force behind evolution, then the more information a system, such as the brain, receives, and the faster it can process this information, the more adequately it will be able to respond to environmental challenges and the better will be its chances of survival. The limit to any intelligent system therefore lies in its abilities to process and integrate large amounts of sensory information and to compare these signals with as many memory states as possible, and all that in a minimum of time. It implies that the functional capacity of a neural structure is inherently limited by its cellular architecture and signal processing time. In other words, there are design principles and operational modes that constrain the potential size and processing power of the brain (see, e.g., HOFMAN 2001b; LAUGHLIN/SEJNOWSKI 2003).

Although the details of the interpretation of the columnar organization of the cortex are still controversial, it is evident that the cerebral cortex is characterized by the hierarchical organization of groups of neurons. To group neurons into functional units interacting over relatively short distances allows these groups to inform as many adjacent clusters of neurons about the state of the 'emitting' cluster with as little as possible redundancy of information. Generally, the growth of connections to units is a factorial function of the number of units in a fully connected network and a linear function of the number of units in a minimally connected network (DEACON 1990; RINGO 1991). Once the brain has grown to a point where the bulk of its mass is in the form of connections, then further increases will be unproductive, due to the declining capability of neuronal integration and increased conduction time. At this point, corresponding to a brain size two to three times that of modern man, the brain reaches its maximal processing power (HOFMAN 2001b). The larger the brain grows beyond this critical size, the less efficient it will become.

Recently, COCHRANE and his colleagues (1998) looked at the different ways in which the brain could evolve to process more information or work more efficiently. They argue that the human brain has (almost) reached the limits of information processing that a neuron-based system allows and that our evolutionary potential is constrained by the delicate balance maintained between conduction speed, pulse width, synaptic processing time, and

neuron density. By modeling the information processing capability per unit time of a human-type brain as a function of interconnectivity and axonal conduction speed they found that the human brain lies about 20-30% below the optimal, with the optimal processing ability corresponding to a brain about twice the current volume. Any further enhancement of human brain power would require a simultaneous improvement of neural organization, signal processing and thermodynamics. Such a scenario, however, is an unrealistic biological option and must be discarded because of the trade-off that exists between these factors. So it seems that within the limits of the existing 'Bauplan' there is no incremental improvement path available to the human brain.

The Problem of Consciousness

The explanation of consciousness is one of the major unsolved problems of modern neuroscience. After several thousand years of speculation, there is at the moment still no agreed answer to the problem of consciousness in explaining how organisms produce representational models of the world, except that most philosophers and neuroscientists do not believe in an immaterial mind which is distinct from the brain. Indeed, it is probable that at any moment some neural processes in the brain correlate with consciousness, while others do not.

Although everyone has a notion of what is meant by being conscious, any attempt at a formal definition of consiousness, however, is likely to be either misleading or overly restrictive. In fact, there are many forms of consciousness, such as those associated with seeing, thinking, pain, emotion and so on. It is plausible that some organisms-in particular the higher encephalized mammals with their complex nervous systems-possess some of the essential features of consciousness, but not necessarily all. It follows that a language system—of the type found in humans-is not essential for consciousness. This is not to say that language does not enrich consciousness considerably, as we shall see later. One might even argue about whether organisms with simple nervous systems, such as insects and molluscs, are not merely cellular reflex automata, but are conscious creatures, who are aware of the external world.

From this perspective, consciousness and affective experience may have arisen concurrently in the evolution of the nervous system; as a way to elaborate and extend the potential reach of instinctual

Evolution and Cognition | 182 | 2003, Vol. 9, No. 2

urges, while new levels of cortical information processing and cognition promoted the ability of organisms to efficiently pursue goals essential to survival. In fact, affective experience, being an intrinsic brain function, can not exist independent of consciousness, since in essence it is something that exists as part and parcel of conscious perception.

In approaching the problem of consciousness, CRICK and KOCH made the tentative assumption that all the different aspects of consciousness employ a basic mechanism or perhaps a few such mechanisms (CRICK/KOCH 1990, 1998). In the case of visual consciousness, for example, they have suggested that its biological usefulness in humans is to produce a single but complex interpretation of the visual scene in the light of past experience, either of ourselves or of our ancestors (embodied in our genes), and to make this interpretation directly available, for a sufficient time, to parts of the brain that make a choice among many different but possible plans of action (CRICK/KOCH 1995). Exactly how this works in detail is unclear.

What is clear is that the term 'consciousness' covers a variety of processes and that some neural elements are essential in constructing a conscious model of the world. But what are these elements and were are they located in the brain. Some form of very short-term memory seems almost essential for consciousness, but this iconic memory may be very transient, lasting for only a fraction of a second, as a kind of 'remembered present' (EDELMAN 1989). Although working memory expands the time frame of consciousness, it is not obvious that it is essential for consciousness (BADDELEY 1992; GOLDMAN-RAKIC 1995). In a similar way, the episodic memory, enabled by the hippocampal system, does not seem to be essential for consciousness (CRICK/KOCH 1998), though an organism without it is severely handicapped. Instead of being restricted to a specific type of neurons or to one part of the 'cortex sytem' it seems more likely that the neural correlate(s) of consciousness are widely distributed over the cerebral cortex and possibly other parts of the brain and involve all types of neurons.

Neural Correlates of Consciousness

To be aware of an object or event, CRICK and KOCH (1995) have argued that the brain has to construct a multilevel, explicit, symbolic interpretation of parts of the visual scene. It means that there are specific groups of neurons at all levels of the visual hierarchy which employ coarse coding to represent some

aspect of the visual scene. In the case of a particular face, all of these neurons can fire to somewhat facelike objects (YOUNG/YAMANE 1992). Notice that while the *information* needed to represent a face is contained in the firing of the ganglion cells in the retina, there is no explicit representation of the face there. A representation of an object or an event will usually consist of representations of many of the relevant aspects of it, and these are likely to be distributed, to some degree, over different parts of the visual system.

The conscious representation of the world is likely to be widely distributed over many areas of the cerebral cortex and possibly over certain subcortical structures as well (BAARS 1997). CRICK and KOCH (1998) postulate that only some types of specific neurons will express the neural correlate(s) of consciousness, and that these neurons will probably be fairly close together and will all project roughly to the same place. An alternative hypothesis is that the neural correlate of consciousness is necessarily global (GREENFIELD 1995). In its most extreme form this would mean that, at one time or another, any neuron in the cortex and associated structures could be part of the neural correlate of consciousness.

The neural correlate of consciousness is defined as the minimal set of neuronal events that gives rise to a specific aspect of a conscious percept (CRICK/ KOCH 2003). The cerebral cortex is probably the most suited part of the brain to look for this neural substrate, as it has very highly and specifically interconnected neuronal networks, many types of excitatory and inhibitory interneurons, and acts by forming transient coalitions of neurons, that is, assemblies of nerve cells, the members of which support one another. The dynamics of coalitions are not simple, as CRICK and KOCH (1990, 2003) have pointed out. In general, at any moment the winning coalition is somewhat sustained and embodies what an animal is conscious of. On the basis of experimental results in the macaque, DESIMONE and DUNCAN (1995) suggest that selective attention biases the competition among competing cell assemblies, but they do not explicitly relate this idea to consciousness.

Coalitions can vary both in size and in character. For example, a coalition produced by visual imagination (with one's eyes closed) may be less widespread than a coalition produced by a vivid and sustained visual input from the environment. These cortical neuronal networks (at least for perception) can be thought of as having nodes. Each node is needed to express one aspect of one percept or an-

Evolution and Cognition | 183 | 2003, Vol. 9, No. 2

other. An aspect cannot become conscious unless there is an essential node for it. For consciousness, there may be other necessary conditions, such as projecting to the frontal cortical areas. Thus a particular coalition is an active network, consisting of the relevant set of interacting nodes that temporarily sustain itself (CRICK/KOCH 2003). The smallest useful node may be a cortical column (MOUNTCASTLE 1998), or perhaps, a portion of a cortical column. The feature which that node represents is (broadly) its columnar property. Recently, EDELMAN and TONONI (2000) presented a theory of consciousness, based on the idea of a 'dynamic core', which resembles the coalition concept to a large extent. The dynamic core hypothesis, however, rejects the idea that there is a special subset of neurons that alone expresses the neural correlate of consciousness, a view which is also defended in the present essay.

Most of the theories of consciousness have the idea of competing assemblies of neurons in common. Consciousness depends on certain coalitions that rest on the properties of very elaborate neuronal networks. It is suggested that attention consists of mechanisms that bias the competition among coalitions, especially during their formation. Furthermore, the idea that the spatio-temporal dimensions of these nodes represent the neural correlates of mind is most appealing, as it suggests that consciousness, being an integral part of the species' problem solving capacity, correlates to some extent with the degree of complexity of a nervous system.

Evolutionary Models of Mind

Considering biological intelligence as the problem solving capacity of an organism makes it possible to speak of degrees of intelligence, and of its evolution from amoeba to man. But what does it mean precisely when one says that species differ in intelligence, or that vertebrates are in general more intelligent than invertebrates? It means that there are differences in the abilities of organisms to perceive and interpret the physical world. Biological intelligence can thus be conceived as to reflect the temporal and spatial complexity of the species' niche, without referring, however, to the kinds of situations organisms encounter in everyday life. It is, in fact, a measure of capacity, independent of the way the capacity is used, and it may be treated as a trait for 'anagenetic' rather than 'cladistic' analysis (GOULD 1976; JERISON 1985). It implies that when distantly related species are comparable in their problem solving capacity, we should consider the species to be comparable in biological intelligence. Yet the near equality in intelligence may be based upon radically different adaptations. Since neural mechanisms and action patterns evolve in the contexts of the environments in which they are effective, and since species never occupy identical niches, many and various intelligences (in the plural) must have evolved in conjunction with evolving environments (JERISON 1985).

In theory, each ecological niche requires its own degree of biological intelligence. That means that specific neural and sensorimotor adaptations, always occur in relation to particular environments. A striking example is the mammalian brain, where the evolutionary changes in the balance of the sensory systems are the result of the adaptive radiation of species into many different ecological niches (PIRLOT 1987; MACPHAIL 1993). These sensory systems, like any other biological feature, could evolve as a result of natural selection, because any subject that forms inadequate representations of outside reality will be doomed by natural selection.

In this view, cognitive systems and emotional phenomena can also be considered to be the result of interactions between genetic aptitude and natural environment, as they have a number of biologically useful functions: one is to keep track of the individual's whereabouts in the world by constructing a schematic model of reality (POPPER 1982; WUKETITS 1986). It is evident that the mind, as an emergent property of sufficiently complex living systems, has its evolutionary history like any other trait that increases adaptation to the environment, and that its functions have increased with the evolution from lower to higher organisms (POPPER 1972).

Evolutionary psychology seeks to explain these evolved, functional characteristics of the human mind through the lense of an explanatory framework where special adaptive modules are postulated to have been critical for hominid survival and reproductive success (see PANKSEPP et al. 2002). These 'adaptive modules' are theoretical constructs unique to the hominid lineage and should be clearly distinguished from the spatio-temporally defined neural processing units (or modules) of the cerebral cortex discussed in the present essay. The existence of a variety of genetically inherited, 'adaptive modules', is dubious at best when considered simultaneously with our current understanding of mammalian brain organization. Indeed, the organization of the cerebral cortex, which is commonly assumed to be a prime anatomical substrate for unique cognitive functions, exhibits no robust

Evolution and Cognition | 184 | 2003, Vol. 9, No. 2

signs of localized anatomical specialization above and beyond specific sensory and motor connections, and their polymodal interactions.

Although adaptation of an organism to its environment is the chief process directing biological evolution, with the evolution of intelligence organisms became more and more independent of their environments, by modifying the environments according to their needs. This process culminated in the evolution of mankind, which can be understood only as a result of the interaction of two kinds of evolution, the biological and the cultural (AYALA 1986). Such considerations have led various authors to argue that the human brain can acquire a large variety of epigenetically derived functions via interactions of a limited number of evolutionary conserved affective/motivational systems (situated largely in subcortical areas) with a set of plastic general-purpose learning mechanisms in the cerebral cortex (see PANKSEPP/PANKSEPP 2000; PANKSEPP et al. 2002). It does not mean that there are no specialpurpose learning systems in the brain, such as fear learning, but the human cerebral cortex includes much more than a conglomeration of special-purpose learning mechanisms. It contains a neural architectecture that can generate flexible features which may be best conceptualized as rewritable.

Cultural evolution, however, being the emergent result of the evolution of mind, cannot dispense with biological preconditions; it builds on biological facts and faculties (VOLLMER 1984; WUKETITS 1986). Though cultural evolution indeed presupposes biological evolution it is not fully explicable in terms of theories and methods of the latter, or as WUKETITS (1986, p199) puts it: "Cultural evolution... has transgressed organic evolution and shows a certain autonomy'. The special status of cultural heredity can be derived from the fact that most cultural innovations are devised precisely in order to meet the environmental challenges or to improve our models of reality, whereas biological evolution has a mindless, random character. It is appropriate therefore, to distinguish adaptations to the environment due to cultural selection from those that take place by the selection of genotypes. Cultural inheritance, furthermore, is an infinitely faster process than genetic inheritance, since it is based on the transmission of information through direct communication, and through books, the arts and the media, which makes that a new scientific discovery or technical achievement can be transmitted to the whole of mankind in less than one generation (Ay-ALA 1986).

Human Language and Intelligence

It is evident that the role of human language in the transmission of knowledge is extremely important, even so prominent and pervasive that it is hardly possible to estimate human general intellectual capacity independent of linguistic capacity (MACPHAIL 1982). Its manifestations and, in particular, that of its newly acquired functions—description and argumentation—is the most peculiar phenomenon in human problem solving. While animals can communicate by expressing their inner state by means of their behavior, and by signaling to conspecifics, (e.g., in case of danger), man is the only creature that is able to make true and false statements, and to produce valid and invalid arguments (see POPPER 1968; POPPER/ECCLES 1977).

The progressive accumulation of interactions between environment (both physical and social), 'conserved' subcortical systems and the 'generalpurpose" cerebral cortex gave rise to a qualitatively different shade of mind—one that could communicate not merely with signs, but in symbolic terms. On the other hand, we have seen that a language system—of the type found in humans—is not essential for consciousness. It is plausible that organisms who do not possess an sophisticated language system, are aware of the external world. This is not to say, however, that language does not enrich consciousness or that it does not contribute to our model of reality.

If we assume that part of the basis of human speech is inherited in the DNA, and that language is as much a biological as a cultural adaptation, then changes in the brain that permit the advantageous supplement of language acquisition to perception and communication would have had obvious selective advantages throughout the period of hominid evolution (DEACON 1998). We may conceive human language, therefore, as a superorganic form of adaptation, evolved not only as a cognitive adaptation contributing to the knowledge of reality of each individual, but also as a means of sharing and, even more importantly, influencing states of mind among conspecifics. Indeed, because of language, human beings are not only able to construct individual representations of the external world, but they can also contribute to and learn from collective models of reality, that is, the cumulative experience of the whole of mankind. With its cognitive and linguistic skills Homo sapiens tries to know its world and even exerts itself to the utmost to control it.

Evolution and Cognition | 185 | 2003, Vol. 9, No. 2

It is obvious that by virtue of language, human beings tend to have highly organized informational states of mind, and, consequently, are excellent problem solvers. But although knowledge of reality may be a necessary condition

for survival it is surely not enough: the degree of intelligence reached by a species does not determine the propensity of its reproductive success. This may be inferred from the indiscriminate elimination of millions of species through the eras, from ammonites to australopithecines. It means that though adaptability increases with the evolution of biological intelligence environmental catastrophes can always be fatal to a species. But not only external factors can threaten the existence of organisms; *Homo sapiens*, despite its impressive intellectual capacities, might in the end become the victim of its own mind by, paradoxically, creating problems that it is then unable to solve.

Concluding remarks

All organisms are constantly engaged in solving problems. Of course, most living creatures are unconscious problem solvers. Even human beings are not always conscious of the problems they are trying to solve. Being aware of these problem situations or not, living organisms must have fitting and relevant models of their specific environments in order to enhance their chances of survival. Consequently, the problem solving capacity of a species is assumed to reflect the temporal and spatial complexity of its ecological niche. The thesis presented here is that biological intelligence can be considered to be a correlate of the problem solving capacity of a species, manifesting itself in the complexity of the species' model of reality.

With the evolution of sensory systems as adaptations to specialized environments, the capacity to process large amounts of sensory information increased and, with that, the power to create more complex physical realities. The processing of large

Author's address

Michel A. Hofman, Netherlands Institute for Brain Research, Meibergdreef 33, 1105 AZ Amsterdam, The Netherlands Email: m.hofman@nih.knaw.nl amounts of information originating from the various sense organs, and the construction of complex models of reality require a neural system that selects, integrates, stores and models: in other words, a system with mind-

like properties that enables the organism to make sense of an otherwise chaotic world. But once we allow mind-like properties to come in, such as motivation, emotion, preference and anticipation, we must allow that it is not only the hostile environment which plays an organizing or designing role in the evolution of biological intelligence, but also the active search of an organism for a new ecological niche, a new mode of living.

Since the mind, prehuman and human, takes a most active part in evolution and especially in its own evolution, hominization and the evolution of our linguistic world may have begun as a cultural adaptation to new ecological niches. The process probably started at the time of hominid divergence a few million years ago, as part of the cognitive and manipulative adaptation to what was in essence a more complex physical reality. In other words, some of the seemingly unique higher functions of the human brain, such as language and other neurosymbolic capacities, were not necessarely due to genetic selection and may have emerged epigenetically through learning and cultural experiences because of the dramatic expansion of the neocortex and its increased tendency to neural plasticity. It seems that the time is finally ripe to begin to building an evolutionary viewpoint of the mind based on comparative concepts that incorporate the intrinsic systems found in all mammalian brains.

Acknowledgements

I am grateful to many of my devoted friends, especially to Michael A. CORNER and the late Herms J. ROMIJN, for their valuable insights and many useful discussions. The comments of three anonymous referees were helpful in improving the manuscript.

References

- Ayala, F. J. (1986) Booknotes. Biology and Philosophy 1: 249–262.
- Baars, B. J. (1997) In the Theater of Consciousness. Oxford University Press: New York.
- Baddeley, A. (1992) Working memory. Science 255: 556–559.
- Ballard, D. H. (1986) Cortical connections and parallel processing: structure and function. Behavioral and Brain Sciences 9: 67–120.
- Buxhoeveden, D. P./Casanova, M. F. (2002a) The minicolumn hypothesis in neuroscience. Brain 125: 935–951.
- **Buxhoeveden, D. P./Casanova, M. F. (2002b)** The minicolumn and the evolution of the brain. Brain Behavior and Evolution 60: 125–151.
- Changizi, M. A./McDannald, M. A./Widders, D. (2002) Scaling of differentiation in networks: Nervous systems, organisms, ant colonies, ecosystems, businesses, universities, cities, electronic circuits, and Legos. Journal of Theoretical Biology 218: 215–237.
- Cherniak, C. (1995) Neural component placement. Trends in Neurosciences 18: 522–527.
- Chklovskii, D. B./Schikorski, T./Stevens, C. F. (2002) Wiring optimization in cortical circuits. Neuron 34: 341–347.
- **Churchland, P. M. (1996)** The Engine of Reason, the Seat of the Soul. MIT Press: Cambridge MA.
- Churchland, P. S./Churchland, P. M. (2002) Neural worlds and real worlds. Nature Reviews/ Neuroscience 3: 903–907.
- Cochrane, P./Winter, C. S./Hardwick, A. (1998) Biological limits to information processing in the human brain. British Telecommunications Publications, http:// www.labs.bt.com/papers/index.htm
- Crick, F./Koch, C. (1990) Towards a neurobiological theory of consciousness. Seminars in Neuroscience 2: 263–275.
- Crick, F./Koch, C. (1995) Are we aware of neural activity in primary visual cortex? Nature 375: 121–123.
- Crick, F./Koch, C. (1998) Consciousness and neuroscience. Cerebral Cortex 8: 97–107.
- Crick, F./Koch, C. (2003) A framework of consciousness. Nature Neuroscience 6: 119–126.
- Deacon, T. W. (1990) Rethinking mammalian brain evolution. American Zoologist 30: 629–705.
- **Deacon, T. W. (1998)** The Symbolic Species: The Co-Evolution of Language and the Brain. Norton and Company: New York.
- **Dennett, D. C. (1983)** Intentional systems in cognitive ethology: The "Panglossian paradigm" defended. Behavioral and Brain Sciences 6: 343–390.
- Desimone, R./Duncan, J. (1995) Neural mechanisms of selective visual attention. Annual Review of Neuroscience 18: 193–222.
- Edelman, G. M. (1989) The Remembered Present: A Biological Theory of Consciousness. Basic Books: New York.
- Edelman, G. M./Tononi, G. A. (2000) A Universe of Consciousness. Basic Books: New York.
- Freeman, W. J. (1975) Mass Action in the Nervous System. Academic Press: New York.
- **Glassman, R. B. (1985)** Parsimony in neural representations: generalization of a model of spatial orientation ability. Physiological Psychology 13: 43–47.
- **Glymour, C. (2001)** The Mind's Arrows: Bayes Nets and Graphical Causal Models in Psychology. MIT Press: Cambridge MA.
- Goldman-Rakic, P. S. (1995) Cellular basis of working memory. Neuron 14: 477–485.
- Gould, S. J. (1976) Grades and clades revisited. In: Masterton,

R. B./Hodos, W./Jerison, H. J. (eds) Evolution, Brain and Behavior: Persistent Problems. Erlbaum: Hillsdale NJ, pp. 115–122.

- Greenfield, S. A. (1995) Journey to the Centers of the Mind. Freeman. New York.
- Hodos, W. (1982) Some perspectives on the evolution of intelligence and the brain. In: Griffin, D. R. (ed) Animal Mind - Human Mind. Springer Verlag: Berlin, pp. 33–56.
- Hodos, W./Campbell, C. B. G. (1990) Evolutionary scales and comparative studies of animal cognition. In: Kesner, R. P./ Olton, D. S. (eds) Neurobiology of Comparative Cognition. Erlbaum: Hillsdale NJ.
- **Hofman, M. A. (1985)** Neuronal correlates of corticalization in mammals: A theory. Journal of theoretical Biology 112: 77–95.
- Hofman, M. A. (1991) The fractal geometry of convoluted brains. Journal für Hirnforschung 32: 103–111.
- Hofman, M. A. (1996) Evolution of the human brain: Design without a designer. In: Elsner, N./Schnizler, H.-U. (eds) Brain and Evolution. Thieme Verlag: Stuttgart, pp. 141– 170.
- Hofman, M. A. (2001a) Evolution and complexity of the human brain: Some organizing principles. In: Roth, G./Wullimann, M. F. (eds) Brain Evolution and Cognition. Wiley and Sons: New York, pp. 501–521.
- Hofman, M. A. (2001b) Brain evolution in hominids: are we at the end of the road. In: Falk, D./Gibson, K. R. (eds) Evolutionary Anatomy of the Primate Cerebral Cortex. Cambridge University Press: Cambridge UK, pp. 113–127.
- Jerison, H. J. (1973) Evolution of the Brain and Intelligence. Academic Press: New York.
- Jerison, H. J. (1985) Animal intelligence as encephalization. Philosophical Transactions of the Royal Society, London B 308: 21–35.
- Jerison, H. J. (1991) Brain size and the evolution of mind. Fifty-ninth James Arthur Lecture on the Evolution of the Human Brain. American Museum of Natural History: New York.
- Klyachko, V. A./Stevens, C. F. (2003) Connectivity optimization and the positioning of cortical areas. Proceedings of the National Academy of Sciences USA 100: 7937–7941.
- Krubitzer, L. (1995) The organization of neocortex in mammals: are species differences really so different? Trends in Neurosciences 18: 408–417.
- Laughlin, S. B./Sejnowski, T. J. (2003) Communication in neural networks. Science 301: 1870–1874.
- Leise, E. M. (1990) Modular construction of nervous systems: A basic principle of design for invertebrates and vertebrates. Brain Research Reviews 15: 1–23.
- Macphail, E. M. (1982) Brain and Intelligence in Vertebrates. Clarendon Press: Oxford.
- Macphail, E. M. (1985) Vertebrate intelligence: the null hypothesis. Philosophical Transactions of the Royal Society, London B 308: 37–51.
- Macphail, E. M. (1993) The Neuroscience of Animal Intelligence: From the Seahare to the Seahorse. Columbia University Press: New York.
- Macphail, E. M./Bolhuis, J. J. (2001) The evolution of intelligence: Adaptive specializations versus general process. Biological Reviews 76: 341–364.
- Malach, R. (1994) Cortical columns as devices for maximizing neuronal diversity. Trends in Neurosciences 17: 101– 104.
- Mayr, E. (1982) The Growth of Biological Thought. Diversity, Evolution and Inheritance. Belknap Press: Cambridge MA.
 Mountcastle, V. B. (1997) The columnar organization of the
- Evolution and Cognition | 187 | 2003, Vol. 9, No. 2

brain. Brain 120: 701–722.

- Mountcastle, V. B. (1998) Perceptual Neuroscience. Harvard University Press: Cambridge MA.
- Panksepp, J./Panksepp, J. B. (2000) The seven sins of evolutionary psychology. Evolution and Cognition 6: 108–131.
- Panksepp, J./Moskal, J. R./Panksepp, J. B./Kroes, R. A. (2002) Comparative approaches in evolutionary psychology: molecular neuroscience meets the mind. Neuroendocrinology Letters, Special Issue, Suppl. 4, 23: 105–115.
- **Pellionisz, A. J. (1988)** Tensor aspects of the multidimensional massively parallel sensorimotor function of neuronal networks. Progress in Brain Research 76: 341–354.
- Pirlot, P. (1987) Contemporary brain morphology in ecological and ethological perspectives. Journal für Hirnforschung 28: 145–211.
- Plotkin, H. C. (1994) Darwin Machines and the Nature of Knowledge. Harvard University Press: Cambridge MA.
- Popper, K.R. (1958) On the status of science and of metaphysics. Ratio 1: 97–115.
- **Popper, K. R. (1968)** Conjectures and Refutations. The Growth of Scientific Knowledge. Routledge and Kegan Paul: London.
- Popper, K. R. (1972) Objective Knowledge. An Evolutionary Approach. Clarendon Press: Oxford.
- Popper, K. R. (1982) The place of mind in nature. In: Elvee, R. Q. (ed) Mind in Nature. Harper and Row: San Francisco, pp. 31–59.
- Popper, K. R./Eccles, J. C. (1977) The Self and its Brain. Springer Verlag: Berlin.
- Prothero, J. W. (1997) Cortical scaling in mammals: A repeating units model. Journal of Brain Research 38: 195–207.
- Rakic, P. (1995) A small step for the cell, a giant step leap for

mankind: A hypothesis of neocortical expansion during evolution. Trends in Neurosciences 18: 383–388.

- **Riedl, R. (1987)** Begriff und Welt: Biologische Grundlagen des Erkennens und Begreifens. Parey Verlag, Berlin.
- **Ringo, J. L. (1991)** Neuronal interconnection as a function of brain size. Brain, Behavior and Evolution 38: 1–6.
- Roth, G./Wullimann, M. F. (eds) (2001) Brain Evolution and Cognition. Wiley and Sons: New York.
- Shettleworth, S. J. (1998) Cognition, Evolution and Behavior. Oxford University Press: New York.
- Szentágothai, J. (1978) The neuron network of the cerebral cortex: A functional interpretation. Proceedings of the Royal Society, London B 201: 219–248.
- **Vollmer, A. (1984)** Mesocosm and objective knowledge. In: Wuketits, F. M. (ed) Concepts and Approaches in Evolutionary Epistomology: Towards an Evolutionary Theory of Knowledge. Reidel, Dordrecht, pp. 69–121.
- Vollmer, A. (1992) The mirror's reflections: Lorenz and evolutionary epistomology. Evolution and Cognition 1: 161– 176.
- Wuketits, F. M. (1986) Evolution as a cognition process: Towards an evolutionary epistemology. Biology and Philosophy 1: 191–206.
- Wuketits, F. M. (1990) Evolutionary Epistemology and Its Implications for Humankind. SUNY Series in Philosophy and Biology: New York.
- Young, M. P./Yamane, S. (1992) Sparse population coding of faces in the inferotemporal cortex. Science 256: 1327– 1331.
- Zeki, S./Bartels, A. (1999) Toward a theory of visual consciousness. Consciousness and Cognition 8: 225–259.

An Evolutionary Model of the Effects of Human Paternal Resemblance on Paternal Investment

A fter making a choice about with whom to mate, females are faced with two potential dilemmas: 1) what is the likelihood that this male is going to invest resources in offspring produced, and 2) does the male provide "good genes,"—that is, offer genetic resiliency to pathogens that will facilitate the offspring's development and reproductive success?

Males. on the other hand, are faced with a different dilemma: making sure that the offspring they invest resources in shares copies of his genes. Males have evolved psychological mechanisms that help to reduce the likelihood of unknowingly raising another man's child. The first mechanism is to attempt to limit his mate's opportunities for extra-pair copulation (EPC) that could eventuate in extrapaternity pair (EPP). Males appear to have evolved three sets of psychological mechanisms designed to reduce, limit,

Abstract

Species with fertilization internal to females face an asymmetry in parental certainty: whereas females are always 100% certain that they will share approximately 50% of their genes in common with their offspring, males can never be certain of paternity. As a consequence of this uncertainty, investment in offspring is similarly asymmetrical and should be, according to evolutionary meta-theory, proportional to the confidence of paternity on the part of males. There are three ways in which a male can attempt to limit cuckoldry. First, he can monitor his mate during the time that she is fertile and engage in tactics that limit her opportunities to engage in extra-pair copulations (EPCs). Second, males have been designed in such a way that if an EPC does occur and results in extra pair paternity (EPP), physiological and behavioral processes out of his conscious control help to limit cuckoldry—namely, sperm competition and semen displacement. Finally, post-parturition he can make assessments about whether the child shares genes in common with him by estimating the degree to which his offspring resemble him. Here I model the third process, the effect that paternal resemblance might have on a male's decisions about how or whether to invest in putative offspring. Additionally, I speculate about the effects resemblance might have in situations of known non-paternity (i.e., step-children and adoption), instances of suspicion about paternity of any one offspring, and the effects of kin selection and kin resemblance.

Key words

Paternal investment, paternal resemblance, social mirror, evolutionary psychology, fluctuating asymme-

and control female infidelity. For example, mate guarding and jealousy during a female's most fertile period will help to reduce and possibly prevent a female's chances of becoming inseminated by anshape of the penis, in particular the glans penis) and behavioral adaptations that capitalize on that morphological adaptation; i.e., the morphology of the human penis allows a male to employ his penis to

Evolution and Cognition | **189** | 2003, Vol. 9, No. 2

other male (Buss 2000). If mate guarding fails to prevent an EPC from eventuating in EPP, males unknowingly employ a number of competitive strategies that help to ensure that his sperm reach the egg, and not another male's. These strategies include, but may not be limited to, sperm competition (BAKER/BELLIS 1995) and semen displacement (GALLUP et al. 2003). Sperm competition entails a number of strategies that enhance placement and volume of sperm in the vagina; e.g., ejaculating small amounts of spermicidal fluids that act to kill an interloper's sperm (BAKER/BELLIS 1995) and increasing semen volume as a function of jealousy and time spent apart (BAKER/BELLIS 1995; POUND et al. 2002), which effectively increases the statistical probability that a male will father the child. Semen displacement is the result of interplay between a morphological adaptation (the

"plunge" the vagina of another male's sperm (see GALLUP et al. 2003). GALLUP et al. (2003, under review) and GOETZ, SHACKELFORD and colleagues (in press) have demonstrated that males behave differently under conditions of possible sperm competition that would maximize the capacity for using their penis to displace another male's semen. Both papers report that under conditions of risk for sperm competition males experienced thrusting deeper, faster, and more vigorously, which would serve as an advantage for semen displacement.

The final defense against cuckoldry is to assess paternity of the child post-parturition. There has been speculation that males do this by assessing the degree to which children share resemblance with him and there is evidence that this strategy is used in humans and perhaps also other species (see e.g., PLATEK et al. 2002, 2003; PLATEK 2002; WESTNEAT/SHERMAN 1997; SHERMAN/MORTON 1988). Phenotypic cues may provide evidence about a man's actual genetic relatedness to his putative offspring. These cues would, of course, concern physical resemblance (Gaulin, personal communication).

BRESSAN and colleagues (BRESSAN/DAL MARTELLO 2002; see also PAGEL 1997) have suggested that males would benefit from his children expressing genes for anonymity of paternity as opposed to paternal resemblance, but this model is not supported by data from several studies (see PLATEK, et al. 2002, PLATEK et al. in press, PLATEK/GALLUP under review). In addition, BRESSAN et al.'s model makes little sense from an evolutionary perspective in which determination of paternity is a critical task for males. Selection for anonymous looking children makes little sense bearing in mind the basic principles of inheritance. The only way in which BRESSAN's anonymity model might be accurate is if the rate of cuckoldry (i.e., EPP) was 50% or greater. However, confirming earlier reports (MACINTYRE/SOOMAN 1991), recent studies using DNA fingerprinting have shown that the incidence of EPP ranges from 1-20% (CERDA-FLORES et al. 1999; SASSE, MÜLLER, CHAKRABORTY/OTT 1994; SYKES/IRVEN 2000). These studies support the idea that there is (and was during human evolutionary history) a substantial risk for EPP as a result of females engaging in EPCs, and males who were selected for detecting EPP would have left more descendants that shared genes in common with them. Following I attempt to model the way in which paternal resemblance might affect a male's decision to invest resources in children.

GEARY (1999) has outlined a theory of paternal investment based on psychosocial–evolutionary strat-

egies. He suggests that paternal investment should only have evolved if it provided some benefit (increases in survival rate of children, quality of child, etc.) to the child (and, by virtue of genetic relatedness, to the parent). Otherwise selection would have favored an abandonment strategy which would allow males the opportunity to engage in additional reproductive opportunities. GEARY emphasizes the importance of paternal presence in the development of offspring, which might have longer-term fitness consequences (see GEARY 1999). For example, children without investing fathers often exhibit subtle social incompetencies, and in the most extreme case may not survive (e.g., because of malnutrition or murder by a resentful stepfather). Therefore, it seems that paternal investment provides relative, not absolute benefit to the child and should result in a mixed reproductive strategy (GEARY 1999). In other words, there should be a large degree of individual variation in investment behaviors as a result of weighing the costs of investment against the costs of losing reproductive opportunities by investing in offspring that may not share genes in common. The costs of investment can be minimized if the male had a mechanism by which he could determine that he is the source of paternity. That is, if the male determines with relative, probabilistic confidence that he is the genetic father, the costs of lost reproductive opportunities are significantly lessened or outweighed if his investment in the child affords him increased fitness with increased long-term fitness advantages; i.e., as a function of the father's investment, the child might develop certain social competencies that afford the child increased reproductive opportunities in the future, enhancing the male's fitness as a consequence.

If the male is uncertain about paternity, then abandonment should have been favored by selection. Additionally, a male could choose to stay with the female (e.g., taking on a wife and step-child) in the interest of mating effort (ANDERSON/KAPLAN/ LANCASTER 1999). Abandonment, however, might allow the male to engage in alternative reproductive opportunities without the cost of investing valuable resources on another man's offspring.

GEARY's model also suggests that selection would have favored investment when there were few alternative reproductive opportunities, which might account for a male taking on a mate and a step-child in the interest of additional mating opportunities . I suggest that this would only happen to the extent that the male had confidence that he would not be at

Evolution and Cognition | **190** | 2003, Vol. 9, No. 2

risk for sperm competition by the female, otherwise the male might be led to abuse, even infanticide, to ensure that his resources were not spent on other male's offspring. Again, this is counter to the argument put forth by BRESSAN and colleagues in which they suggest that males would experience increased fitness returns by not knowing paternity. This would only be the case for those males chosen as EPC partners. One could hypothesize that there is something about males who engage in EPCs, who are selected as EPC partners, or who adopt a primarily EPC-driven reproductive strategy to produce anonymous looking children. However, because the rates of EPP are on the order of 1–20%, the majority of males in the population would still require a means by which to make assessment about paternity. It therefore is probably more parsimonious to presume that females choose EPC partners along two dimensions that help to ensure investment from their in-pair partner. First, females could choose to engage in EPCs with males who resemble their in-pair mate, which would increase the likelihood that the child might also bear some resemblance to the in-pair mate. This would also aid in maternal warping of the social mirror effect (Gaulin, in press); i.e., if a female chooses an EPC partner who resembles her in-pair partner, the in-pair partner may be more susceptible to social affirmations of resemblance. Second, because females usually choose EPC partners that are more attractive, symmetrical, and are presumed to have "better genes" (see BAKER/BELLIS 1995, for a review), an inpair mate might be more inclined to react favorably towards a child that expresses "good genes" (e.g., being more symmetrical). MEALEY (2003) recently demonstrated that symmetry among twins did not account for variation in treatment of the twins. However, PLATEK et al. (in preparation), using a morphing paradigm have documented significant differences in reactions to children's faces as a function of symmetry, with more symmetrical faces eliciting greater positive investment responses. Although the effects of symmetry may play an important role in child investment strategies, symmetry is not modeled here.

Evidence for the Effects of Genetic Relatedness and Paternal Resemblance

There is evidence that males differentially invest in children to whom they are genetically related. It is not uncommon for unrelated, or otherwise stepchildren to be treated significantly worse than biological children (e.g., ANDERSON et al. 1999). BURCH and GALLUP (2000) have shown that the less a male thinks a child looks like him, the worse he treats the child, and the worse the relationship with that child is. DALY and WILSON (1983; DALY/WILSON/WEG-HORST 1982) have shown that male sexual jealousy can become violent and can be interpreted as a byproduct of the fear of infidelity, or possibility of becoming a cuckold. They estimate the incidence of abuse that results in infanticide among step-children to be 100 times that of genetically related children. These data suggest a strong link between perceived resemblance, relatedness and the treatment of children.

However, in terms of assessing resemblance, the absence of mirrors during human evolution might have precluded males from being able to objectively determine the degree of facial resemblance. There are several hypotheses about how and whether resemblance could have had effects during our evolutionary history. The first falls under a category that has been referred to as the social mirror effect (BURCH/GALLUP 2000). A social mirror consists of 1) other people telling you that a child resembles you, 2) making comparisons between your child and your putative genetic kin, and 3) monitoring behavioral similarities, or lack thereof, between your offspring, your kin, and yourself. There are data to support the existence of a social mirror. It has been shown that other's perceptions of offspring-paternal resemblance is important in 1) the father's perceptions of resemblance (DALY/WILSON 1983; REGAL-SKI/GAULIN 1993) and 2) the degree of violent behavior enacted by the father toward the offspring and the mate (BURCH/GALLUP 2000, in press). Further, DEBRUINE (2002) has shown that people trust those who resemble them more than those who do not, suggesting a role for kin-resemblance comparisons (see also PLATEK et al. 2003).

However, although social mirror information seems to be important for the ways in which a male behaves toward a child, my colleagues and I have recently found that, in hypothetical investment situations, males react more positively toward faces of children that actually share more of their own facial features (PLATEK et al. 2002). This morphing study employed stimulus faces that represented a 50% morph; i.e., the stimulus child's face was 50% child and 50% subject. We have found that, in order for a male to change from reacting indifferently to favorably toward a child's face they must share at least 25% facial characteristics in common with the child's face (PLATEK et al. 2003). This proportion of shared characteristics parallels the degree of shared

Evolution and Cognition | 191 | 2003, Vol. 9, No. 2

genes in relatives two-steps removed (e.g., nieces, nephews, and grandchildren), which provides indirect support for the notion that, during human evolution investment decisions not only in a male's own children, but also in his kin's children and grandchildren, may have been dependent upon a resemblance perception mechanism/module.

Some other hypotheses about how our ancestors might have assessed resemblance include the water reflection hypothesis; i.e., our ancestors might have been able to see themselves in reflections on the water's surface (Gallup, personal communication)-in fact chimpanzees have been observed investigating their images in pools of water and urine (KEENAN, GALLUP/FALK 2003). This hypothesis, although intriguing, is highly unlikely for three reasons: first, rarely is water so still that one can stare at it and gather accurate information from one's reflection and, moreover, the probability of still waters occurring on a regular basis was probably also rare. Second, a reflection of water is skewed by any motion in the water and by the distortion due to murkiness. Third, an adequate account of the reasoning for why our ancestors would have benefited from contemplating their existence while staring in the water has yet to be developed. Further, it has been postulated the possibility that the water reflection hypothesis was a driving force behind the selection against mirror self-recognition among gorillas, stating that gorillas who stayed by the water side long enough to contemplate their own existence were at greater risk of predation by water-dwelling predators such as crocodilians and reptiles.

Another hypothesis has to deal with proprioceptive responses and suggests that an individual might get information about the way he or she looks by 1) touching his or her own face and 2) by internal visceral proprioceptive responses from the underside of the face. Both seem plausible in light of the fact that the somatosensory system is highly connected to the visual cortex and that the human brain creates visual maps associated with tactile information (e.g., MACALUSO/FRITH/DRIVER 2000; see also HAUBER/SHERMAN 2001) for discussions of self-referent phenotype matching). Although these alternative hypotheses are tenable, recent neuroimaging (fMRI) data suggests the presence of fundamental neurocognitive sex differences in reactions to children's faces as a function of resemblance (PLATEK et al., under review). It is likely that resemblance was assessed via a number of sensory channels in the ancestral environment and may incorporate a number of the hypotheses that have been generated. Additionally however, PLATEK, et al. (under review) have recently demonstrated sex differences in neural activity as a function of children's faces based on resemblance. Unlike females, who showed activation typical of mental state attribution, males demonstrated activity in the left frontal lobes in areas associated with inhibition on negative responding. These findings suggest that males may posses a generalized skepticism about children that is alleviated when they detect resemblance.

Paternal investment is likely affected by variables other than resemblance as well. For example, birth order has also been suggested as being related to the degree and amount of abusive acts by the father (DALY/WILSON 1984). An individual born early in the birth order is less likely to incur abuse at the hand of the father. This phenomenon should be predictable to the extent that each additional child represents 1) an instance of male ambiguity regarding the paternity of that particular child, 2) the effects that ambiguity about paternity has on his confidence of the paternity of his other children, and 3) an increased draw on his economic resources. Every time a male's mate gives birth to a child, the likelihood that the child shares genes in common with him is approximately 50% (i.e., the child is either his or not, but males are probably more certainty because the rate of EPP is estimated at 1-20% the likelihood that a child shares genes in common with him might be more like 80-99%). For example, each additional child means increased psychological strain for the male in the form of mate guarding, suspicion of infidelity, physiological and behavioral energy in the form of sperm competition and semen displacement, allocation of resources and also having to assess resemblance again. Furthermore, if the male determines that subsequent children do not share genes in common with him, he may begin to question the paternity of his other children and engage in abusive behaviors or ultimately abandon the family in order to pursue alternative reproductive opportunities.

Finally, resource allocation as a function of number of offspring may related to offspring viability. The more resources a parent invests in a child, the higher the probability of that child reaching sexual maturity. BECKER (1991) and DALY/WILSON (1984) theorize that parents can maximize their investment benefits by discriminate investment according to age of offspring and quality of offspring. They suggest that, under situations of environmental depravity it would be most beneficial to invest in older offspring because they have the highest probability

Evolution and Cognition | **192** | 2003, Vol. 9, No. 2

of surviving to reproduction given that they have survived as long as they have and the time left to reach reproductive maturity is shorter than it is for younger children. However, in varying environmental conditions parental investment in the form of resource allocation would co-vary as a function of the changing environment (DAVIS/TODD 1999). In this model I assume that environmental situations are relatively stable, and the simple, "fast and frugal" environmental decision making is left to relatively equal allotment of resources among offspring; i.e., with a constant amount of resources to allocate, each additional child represents an equal draw on that resource pool, thus lessening the amount for each child proportionally in multiple child families (DAVIS/TODD 1999; for a more in depth description of environmental decision making and parental investment, see GIGERENZER/TODD 1999). In addition, I would like to expand with a sentence on Becker's hypothesis to include the idea that parents might also differentially invest in those offspring that they feel have the best chance of surviving; i.e., by assessing some trait(s) that might bear on survival and reproduction such as fluctuating asymmetry. Males who differentially invest in offspring or their kin's offspring that are more symmetrical are increasing their chances for positive fitness consequences in the long run.

This theoretical paper attempts to model some of the parameters that contribute to paternal investment using data on child abuse and spousal abuse (BURCH/GALLUP 2000; BURCH/GALLUP under review) and experimental data that suggests that the degree to which a child shares facial features with a male affects the male's decisions to allocate resources more than it affects females' (PLATEK et al. 2002, 2003; PLATEK 2002). The model only aims to explain the effects of men's reactions towards children as a part of a broader model of human offspring viability. I will attempt to model the interaction between a male's assessment of paternal resemblance and the number of children he has as a function of the probability that any one child is the result of a female EPC. I propose two possible models: an additive model which assumes that social mirror-mediated resemblance and actual resemblance summate in their effects on male investment strategies and a deception detection model which assumes that females employ social mirror-mediated resemblance as a way to cuckold their mates. I also attempt to cast all possible aspects of the model into clearly testable predictions.

Paternal Investment as a Function of Paternal Resemblance: An Additive Model of Resemblance

According to the models presented in this paper, paternal resemblance can be seen as two separate, but related phenomena (see BREDART et al. 1999; BURCH/ GALLUP 2000; MCCLAIN et al. 2001; NESSE/SILVER-MAN/BORTZ et al. 1990; PLATEK et al. 2002, under review). One aspect of paternal resemblance is that any individual offspring accurately, or to some degree, physically resembles the father, as shown by CHRISTENFELD/HILL (1995; but see NESSE/SILVERMAN/ BORTZ 1990; BREDART et al. 1999; MCCLAIN et al. 2000) and as suggested by the principles of inheritance. This would represent an instance of observable shared phenotypic traits; i.e., alleles shared by descent. This process is probably occurring unconsciously, see PLATEK et al. (2002, 2003, under review; PLATEK 2002). The other, and equally important, aspect of paternal resemblance is the degree to which one's own and other's perceptions of the offspring's resemblance to the father are taken into account: the social mirror (BURCH/GALLUP 2000; DALY/WIL-SON 1982; REGALSKI/GAULIN 1993). The interaction between actual resemblance and social mirror-mediated resemblance can be represented as:

Paternal resemblance =
$$(P_{rt})(P_{rs}) = P$$
 (1)

where P_{rt} represents the proportion of shared resembling traits or the actual degree of paternal resemblance, and P_{rs} represents the proportion of times others "socially" assert paternal resemblance: social mirror-mediated resemblance. The product of each of these proportions will be referred to as *P* measured on a scale of 0 to 1, with 0 indicating neither type of paternal resemblance being present and 1 indicating that the most possible of each type is present.

Paternal investment, however, depends on several factors. An increase in paternal investment can increase the likelihood of offspring survival to sexual maturity and social competency, ultimately having effects on a child's fitness, thus the mechanisms driving the behaviors are likely to be complex. Therefore, paternal investment is modeled here to be represented by P divided by the probability that any individual offspring could be the result of an EPC or .50; i.e., the child is either the offspring of the male in the relationship or some other male, thus leaving roughly a 50:50 probability that a male decides that any offspring was or was not fathered by him. This is referred to as the probability of ge-



Figure 1: Graph depicting the additive and deception predictions from this model. If there is no actual resemblance and no one ascribes resemblance of the child to the dad then male parental investment should be low, if present. If the actual resemblance is present, but social mirror-mediated resemblance is not *or* actual resemblance is not present, but social mirror-mediated resemblance is present than male parental investment should be somewhere in the middle probably with much between male variance; corresponds to a coefficient of abuse score of 1. The optimal condition for male parental investment would be a situation in which both actual and social mirror-mediated resemblance is present. Females would be predicted to invest equally among all conditions since maternal resemblance in this model is not predicted to affect female behavior towards children. Filled circles represent female reactions under the additive model; X-filled circles represent male reactions under the deception model.

netic non-relatedness, *G*. It is important again to keep in mind that probability of cuckoldry probably varies dramatically from 50% and is probably on the order of 1–20% (CERDA-FORES et al. 1999; SASSE, MULLER, CHAKRABORTY/OTT 1994; SYKES/IRVEN 2000). Therefore, if *G* is multiplied by an estimate of the cuckoldry rate (C_k) we can obtain an estimate of a male's relative certainty of paternity as a function of an estimate of the population estimates of EPP that will be referred to as *C*:

Therefore,
$$G \cdot C_k = C$$
 (2)

The model equation for paternal investment (*PatI*) can be reduced to:

$$PatI = P/C \tag{3}$$

The equation might be viewed as a *coefficient of paternal investment*, given that as paternal certainty decreases as a function of decreased perceived or actual paternal resemblance, the likelihood that the male would be motivated to engage in offspring abuse would be expected to increase and investment would be expected to decrease. Furthermore, holding *P* constant and changing the relative rate of EPP will affect male investment in a manner consis-

tent with contemporary evolutionary meta-theory. *P* is the product of P_{rt} and P_{rs} , which are proportions, and thus must be less than 1, with 1 being representative of the highest degree of actual paternal resemblance and the highest degree of social affirmation of resemblance (both abstract and difficult to measure). If we substitute modern estimates of EPP (range 1–20%), we find that an optimal paternal investment score ($P_{rt} = 1$, $P_{rs} = 1$, $C_k = 0.5$, and G = 0.01) would equal 400 and decrease logarithmically to the lowest possible degree of paternal investment (or depending upon the interpretation paternal abuse) to 0.1 ($P_{rt} = 0.1$, $P_{rs} = 0.1$, $C_k = 0.5$, G = 0.2).

My colleagues and I have collected data pertaining to the paternal resemblance portion of this model. We have found that males bias their reactions to children's faces, and hypothetically invest more in children that share more facial characteristics with them (PLATEK et al. 2002, 2003, under review; PLATEK 2002). We have also begun to test how actual or social mirror-mediated resemblance accounts for the variance in the way a male reacts towards the faces of children. The model predicts that when both actual and social mirror-mediated resemblance are present, a man's reactions toward that child should be overwhelmingly positive, whereas when neither are present his reac-

Evolution and Cognition | 194 | 2003, Vol. 9, No. 2

tion should be overwhelmingly negative. When the two differ in relative proportions, an additive model would predict negation, which might result in indifference toward a child. Four different, discrete conditions could arise and were tested in a follow-up study: (1) a child shares no resemblance with dad and people do not ascribe resemblance of the child to dad, (2) a child shares no resemblance with dad, but people ascribe resemblance of the child to dad, (3) a child shares resemblance with dad, but people do not ascribe resemblance of the child to dad, and (4) a child shares resemblance with dad and people also ascribe resemblance of the child to dad. When predicting the behaviors towards the child and the viability of the child in terms of paternal investment, the last option would be the optimal condition and the first would be the worst condition if the assumption of additivity were adopted (see Figure 1 for a graphical depiction of the predictions). PLATEK (2002) tested an inbetween teraction social mirror-mediated information and actual shared resemblance and found that males and females are affected similarly by social mirror information, but only males appear to be impacted by shared resemblance.

An exponent term, 1/i, could be added to the equation to represent the number of offspring an individual has. As can be seen, increases in the number of offspring is predicted to yield decreases in the amount of investment, or increases in abuse by the male, because the probability that he has been cuckolded has increased, because each additional child that the mother bears represents an incidence in which an EPC could have resulted in EPP (BAKER/ BELLIS 1995; SCHACT/GERSHOWITZ 1963). With one child the probability of cuckoldry modeled here is GC_k . With each additional child the male faces the possibility that he was cuckolded again; this would account for increases in the amount of abuse seen with each subsequent child in the birth order theorized here to be the result of increased psychological strain on mechanisms designed to assess resemblance and predict probability of cuckoldry. Abuse related to birth order effects has been demonstrated in humans (BURCH/GALLUP unpublished data; DALY/WILSON 1982; REGALSKI/ GAULIN 1993). Further, additional children increase the strain on that male's resources. Additional children represent increased costs; i.e., if he is now expected to invest more (e.g., time, resources, etc.) in offspring, this might occur at the cost of lost additional mating opportunities.

$$PatI = P^{1/i}/C \tag{4}$$

An Alternative Model: Deception Detection and the Social Mirror

The paternal resemblance component in the above equation assumes additive effects of actual and social mirror resemblance; i.e., the above model assumes that social mirror-mediated and actual paternal resemblance sum to produce an effect that is the probability of paternal investment. However, this may not be the case. Females may try to manipulate social mirror information to deceive males in to believing that an offspring shares genes in common with him and thus invest in offspring fathered by another male (DALY/WILSON 1982; GAULIN, personal communication; REGALSKI/GAULIN 1993). It would have behooved females to assert paternal resemblance if perceptions of resemblance translated to increased investment on the part of the male and that investment resulted in greater offspring survival. However, a male who blindly trusted a female's assertions may have been at risk to invest in offspring to whom he was genetically unrelated. As a consequence of social mirror-mediated resemblance, perhaps being an instance of deception on the part of the female, it would have been in males' best interests to evolve a deception detecting mechanism; i.e., to question paternity when females enthusiastically assert paternity. This counter model can be modeled by subtly altering the original equation:

$$PatI = [(P_{rt})(IP_{rs} \log_{10} I)^{1/i}/C] = P'$$
(5)

thus,

$$PatI = P'/C \tag{6}$$

By representing social mirror-mediated resemblance as the absolute value of the base 10 log of the observed P_{rs} , the return decreases the weight of the paternal resemblance term of the equation, thus lessening a male's investment. When social mirrormediated assertions of resemblance ensue this effect is likely impacted by who is providing the information. For example, it might be the case that, when social mirror assertions are provided by the maternal relatives, males might place lesser weight on that information, and may even weight that information negatively. However, when social affirmations of resemblance are provided by the male's kin, positive weights might be applied and the social mirror effect might have a greater impact on paternal investment (see below).

Evolution and Cognition | **195** | 2003, Vol. 9, No. 2

Accounting for Genetically Unrelated Children: Step- and Adopted Children

A slight modification of this model can also be used to account for how a male might invest resources in children to whom he is certain to be genetically unrelated, such as step-children, foster and adopted children. If a child is unrelated to the male, then there is no need to take into account the denominator term (C) of the equation presented in equations (3) and (6) above. In other words, if resemblance affects males as we predict it does, then investment in unrelated children ought to be directly related to the degree to which unrelated children also resemble the male. Therefore,

$$PatI_{(\text{UnRel})} = P^{1/i}/1 \tag{7}$$

This equation takes into account only the effects that resemblance has on the male's reactions toward the child, which may translate into investment strategies as we have shown in our behavioral studies (e.g., PLATEK et al. 2002), because he will be 100% certain that the children he is caring for share no genes in common with him. This modification of the model may be less important for understanding human evolution than it might be for understanding how males react to non-genetically related children and may provide insight for social welfare practice in placing children with foster or adoptive parents.

Accounting for Suspicion of Cuckoldry for Any One Child

If a male suspects that any one of his children are the result of his partner having cuckolded him, then this ought to raise uncertainties about the paternity of his other children; i.e., his partners general faithfulness toward him. Thus, the degree of genetic relatedness term (*G*), but not C_k , is affected by the number of children he has if he is concerned that one of those children were the product of an EPC, which will increase the theoretical threshold for investing resources in any of his other children, affected by an increased or reinstatement of a generalized skepticism about paternity. Such a male may need to re-assess resemblance to all of his children, something that he would have otherwise not had to do.

$$PatI_{(Sus 1)} = P/G^{1/i}C_k \tag{8}$$

Thus, the probability of paternal investment when the male suspects one child is the byproduct of cuckoldry

$$(PatI_{(Sus 1)}) = P/C' \text{ or } P'/C'$$
(9)

where C' is equal to the denominator term of equation (8). Equation (9) increases the threshold with which the level of paternal investment needed to reach a level equivalent to that if there was no suspicion of cuckoldry about any one of his children.

Kin Selection and the Social Mirror Effect

The effects of the social mirror may be mediated in part or in whole by the individual(s) who provide the information to the male (e.g., is this person a genetic relative or not?). In other words, kin selection may play a role in the way males react toward social mirror information. With this in mind, the model may be modified to account for this by altering P_{rs} . P_{rs} becomes the mean proportion of time that someone says a child resemble the putative father coded for the degree of genetic relatedness of the person providing the information. People who share 50% or more of their genes in common with the male are coded as 1. Those sharing 25% of their genes in common with the male are coded as .5. Those sharing .25% of their genes in common with the male are coded as .125, and so on. Thus the less someone is genetically related to the male, the less credence their social mirror information should have on influencing the male's decisions about whether to invest or not. This mean or proportion codes becomes P_{rs} in equations (3, 6, 7, or 9). According to this modification, when someone shares no genes in common with the male, but attempts to assure resemblance using the social mirror the degree of paternal investment should drop to zero, which is similar to the deception model (equation 6) presented above.

Concluding Comments

Clearly, it is over-simplistic to think that paternal resemblance and the probability of cuckoldry are the only forces that affect paternal investment. As well, it is likely that there are as intricate dimensions to the amount of maternal investment allocated that have yet to be elucidated clearly. However, paternal certainty clearly plays an important role in the in-

Evolution and Cognition | 196 | 2003, Vol. 9, No. 2

Author's address

Steven M. Platek, Department of Psychology,

Drexel University, 245 N. 15th St., Mail Stop

626, 10th Floor Bellet Bldg., Philadelphia,

Email: steven.m.platek@drexel.edu

vestment of children and, therefore, models that help clarify the mechanisms involved in a male making a determination about paternity and investment strategy will aid in understanding investment decisions and paternal abuse. This is a theoretical pa-

per, a working hypothesis of the ways in which these phenomena *might* be having effects on human offspring viability and in no way suggest these

PA 19102

are the only mechanisms that are responsible in this process. However, each of the hypotheses is presented in clearly testable, empirically falsifiable ways.

Acknowledgements

The author thanks Gordon GALLUP and Todd SHACK-ELFORD for comments on an earlier version of this manuscript.

References

- **Bressan P./Dal Martello M. F. (2002)** Talis pater, talis filius: Perceived resemblance and the belief in genetic relatedness. Psychol Sci. 13: 213–218.
- Burch, R. L./Gallup, G. G. Jr. (2000) Perceptions of paternal resemblance predict domestic violence. Evolution and Human Behavior 21: 429–435.
- Burch, R. L./Gallup, G. G. Jr. (under review) Domestic violence as a stimulus for pregnancy. Journal of Family Violence.
- Buss, D. M. (2000) The Dangerous Passion: Why Jealousy Is as Necessary as Sex and Love. Free Press: New York.
- Cerda-Flores, R. M./Barton, S. A./Marty-Gonzales, L. F./Rivas, F./ Chakraborty, R. (1999) Estimation of nonpaternity in the Mexican population of Nuevo Leon: A validation study with blood group markers. American Journal of Physical Anthropology 109: 281–293.
- Christenfeld, N. J. S./Hill, E. A. (1995) Whose baby are you? Nature 378: 669.
- Daly, M./Wilson, M. (1983) Sex, Evolution, and Behavior. Second Edition. Willard Grant: Boston.
- Daly, M./Wilson, M. (1982) Whom are new born babies said to resemble? Ethology and Sociobiology 3: 69–78.
- Daly, M./Wilson, M./Weghorst, S. (1983) Male sexual jealousy. Ethology and Sociobiology 3: 11–27.
- Davis, J. N./Todd, P. M. (1999) Parental investment by simple decision rules, In: Gigerenzer, G./Todd, P. M. (eds) Simple Heuristics that Make Us Smart. Oxford University Press: New York.
- Geary, D. C. (2000) Evolution and proximate expression of human paternal investment. Psychological Bulletin 126: 55–77.
- Gigerenzer, G./Todd, P. M. (1999) Simple Heuristics that Make Us Smart. Oxford University Press: New York.
- Hauber, M. E./Sherman, P. W. (2001) Self-referent phenotype matching: Theoretical considerations and empirical evidence. Trends in Neuroscience 24: 609–616.
- MacIntyre, S./Sooman, A. (1991) Non-paternity and prenatal genetic screening. Lancet 338: 869–871.
- Macaluso, E./Frith, C. D./Driver, J. (2000) Modulation of Human visual cortex by cross modal spatial attention. Science 289: 1206–1208.

- Mealey, L. (2003) Paper presented at annual meeting of the Human Behavior and Evolution Society, Lincoln NE.
- Nesse, R./Silverman, A./Bortz, A. (1990) Sex differences in ability to recognize family resemblance. Ethology and Sociobiology 11: 11–21.
- **Platek, S. M. (2002)** Unconscious reaction to children's faces: The effect of resemblance. Evolution and Cognition 8: 207–214.
- Platek, S. M./Burch, R. L./Panyavin, I. S./Wasserman, B. H./ Gallup, G. G. Jr. (2002) Reaction to children's faces: Resemblance matters more for males than for females. Evolution and Human Behavior 23 :159–166.
- Platek, S. M./Critton, S. R./Burch, R. L./Frederick, D. A./Myers, T. E./Gallup, G. G. Jr. (2003) How much resemblance is enough? Sex differences in reactions toward children's faces, but not in detection of resemblance. Evolution and Human Behavior 24: 81–87.
- Platek, S. M./Davis, J. A./Mohamed, F. B./Gallup, G. G. Jr. (under review) Children on the mind: Sex differences in neural correlates of facial resemblance. Evolution and Human Behavior.
- **Regalski, J. M./Gaulin, S. J. C. (1993)** Whom are Mexican newborns said to resemble? Monitoring and fostering paternal confidence in the Yucatan. Ethology and Sociobiology 14: 97–113.
- Sasse, G./Muller, H./Chakraborty, R./Ott, J. (1994) Estimating the frequency of non-paternity in Switzerland. Human Heredity 44: 337–343.
- Schact, L. E./Gershowitz, H. (1963) Frequency of extra-marital children as determined by bloodgroups. In: Proceedings of the Second International Congress in Human Genetics, pp. 894–897.
- Sherman, P. W./Morton, M. L. (1988) Extra-pair fertilizations in mountain white-crowned sparrows. Behavioral Ecology and Sociobiology 22: 413–420.
- Sykes, B./Irven, C. (2000) Surnames and the Y chromosome. American Journal of Human Genetics 66: 1417–1419.
- Westneat, D. F./Sherman, P. W. (1997) Density and extrapair fertilizations in birds: A comparative analysis. Behavioral Ecology and Sociobiology 41: 205–215.
- Wilson, E. O. (2000) Sociobiology: The New Synthesis. 25th Anniversary Edition. Presidents and Fellows of Harvard College. Harvard University Press: Cambridge.

Evolution and Cognition | **197** | 2003, Vol. 9, No. 2

Zusammenfassungen der Artikel in deutscher Sprache

Arthur S. Reber Einige überraschende Konsequenzen der kognitiven Revolution

In dieser Arbeit wird ein Überblick über die sog. "kognitive Revolution", die in der Mitte des letzten Jahrhunderts begann und nach wie vor mit beträchtlichem Enthusiasmus fortgesetzt wird vermittelt. Dabei werden vor allem die Übergänge einer behavioristisch dominierten Denkweise hin zu jener Psychologie untersucht, die subjektive Zustände, kognitive Funktionen und darunter liegende neuronale Strukturen in Betracht zieht. Was diese Revolution jedoch übersehen hat ist der Vorgang des Lernens und Anpassungsphänomene – wie sie ein darwinistische Perspektive nahe legt.

Das Thema "Lernen" wurde dabei vielfach durch Untersuchungen von Gedächtnis und Repräsentation ersetzt, während der Prozess der eigentliche Anneigung von Wissen zu wenig Beachtung findet.

Ebenso lassen viele kognitivistischen Modelle folgende heuristische Funktionen Darwinischen Denkens ausser Acht: hat das untersuchte System adaptiven Wert? Konnte es sich innerhalb entsprechender Zeiträume auch entwickeln?

Manfred Hassebrauck Die Auswirkung des Fertilitätsrisikos auf die Beziehungseinschätzung

In dieser Studie wurde der Einfluss des weiblichen Menstruationszyklus auf die Verarbeitung von Informationen über die eigene Paarbeziehung untersucht. 87 Frauen machten zunächst globale Angaben über ihre Zufriedenheit mit der eigenen Beziehung und bewerteten diese dann im Hinblick auf 64 spezifische Merkmale, die den Prototyp einer guten Paarbeziehung (Hassebrauck/Fehr 2002) charakterisieren. Frauen, die keine hormonellen Kontrazeptive benutzten, verarbeiteten während der fertilen Phase ihres Zyklus Informationen über ihre Beziehung systematischer als Frauen während der nicht fertilen Phase. Die Ergebnisse werden in einem evolutionären Kontext diskutiert.

Steve Stewart-Williams

Über den möglichen evolutionären Ursprung des Glaubens an die Existenz einer externen, bewusstseinsunabhängig existierenden Aussenwelt

In dieser Arbeit steht der sog. "metaphysische Realismus" und dessen mögliche evolutionäre Begründung im Zentrum. Unter "metaphysischem Realismus" versteht man die tief verwurzelte Neigung des Menschen an eine bewusstseinsunabhängig existierende Aussenwelt zu glauben. Eine evolutionäre Begründung dieser Tendenz geht einmal davon aus, dass der menschliche Geist über die Fähigkeit verfügt mentale Zustände mit "objektiven Bezugspunkten" von solchen zu unterscheiden, die keine derartigen Bezugspunkte aufweisen. Weiters verfügt der menschliche Geist über die Fähigkeit der dauerhaften Repräsentation von nicht unmittelbar wahrgenommenen Aspekten der Welt.

Diese Fähigkeiten werden als evolutionäre Strategien interpretiert, welche stark fitness steigernde Wirkung haben. So würde beispielsweise die Unterscheidung "subjektiv" versus "objektiv" sich darauf beziehen, dass unmittelbare Wahrnehmung andere Verhaltenskonsequenzen erfordert als mentale Vorstellungsaktivität.

Ausgehend von diesen Überlegungen liegt die Annahme einer genetisch–evolutionären Bedingtheiten des "hypothetischen Realismus" nahe.

O'Connell/R. Dunbar Ein Test für das Erfassen von "false belief" bei Schimpansen

"False belief" Aufgaben werden im Bereich der Entwicklungspsychologie dazu verwendet eine "theory of mind" bei Kindern nachzuweisen. Bei Tierversuchen wurde ähnliche Kompetenzen im Bereich der sog. "sozialen Kognition" untersucht, welche als Vorstufen einer "theory of mind" angesehen werden. Jedoch wurden eigentlich "false belief" Tests bei Tieren nicht durchgeführt.

In dieser Arbeit werden die Ergebnisse eines nonverbalen, mechanischen – einem "false belief" Test analogem Test vorgestellt. Getestet wurden 4 Schimpansen, 11 autistische Erwachsene und 41 Kinder im Alter zwischen 3 und 6 Jahren. Die Ergebnisse legen die Annahme nahe, dass die mentalen Fähigkeiten von Schimpansen denen von Kindern vor der Entwicklung einer "theory of mind" ähnlich sind.

Fred Heeren Frühe Craniata und der Weg zur Kognition?

Haikouella lanceolata erscheint derzeit als das älteste, gut dokumentierte Chordatier. Dieses Tier besitzt ein relativ grosses Gehirn und wirft damit die Frage auf, ob Gehirn und Endoskelett gleichzeitig evoluierten - wie bisher angenommen wurde, oder ob das Gehirn lange Zeit vor der Ausbildung eines Endoskelettes vorhanden war. Paläontologische Untersuchungen dazu sprechen von einem "top down" Muster im Auftreten neuer Formen im Fossilbestand. Davon ausgehend stellt sich auch die Frage ob menschliche Kognition gleichsam unvermeidlich im Verlauf der Evolution auftraft, oder ob diese eher durch bloße historische Zufälligkeiten entstand. Verläuft Evolution "bottom up" indem kleine Veränderungen unter natürlicher Selektion zu größeren Veränderungen führen, oder sind hier auch "top down" Kräfte wirksam, welche die Entwicklung von Organismen nur innerhalb bestimmter Formen zulässt?

Ingo Brigandt Gestalt-Experimente und induktive Beobachtungen. Konrad Lorenz' frühe erkenntnistheoretische Schriften und die Methoden der klassischen Ethologie

Konrad Lorenz formulierte seine frühen erkenntnistheoretischen Auffassungen während der vierziger Jahre, wobei er die Bedeutung der kognitiven Prozesse, der Induktion sowie der Gestaltwahrnehmung betont. Nach dem zweiten Weltkrieg benutzte er diese philosophische Theorie, um die Methodik der klassischen Ethologie gegenüber anderen Ansätzen in der Verhaltensforschung zu verteidigen. Dieser Beitrag diskutiert das Verhältnis zwischen Lorenz' ethologischer Methodologie und seiner Erkenntnis- und Wissenschaftstheorie.Dabei wird deutlich, dass Lorenz deutliche Parallelen zwischen beiden zieht. Auf der methodologischen Ebene kommt für Lorenz die Beobachtung logisch und zeitlich vor dem Experiment. Erkenntnistheoretisch hat die Gestaltwahrnehmung Vorrang vor der Induktion. Während die Beobachtung von Verhaltensmustern hauptsächlich vom kognitiven Prozess der Gestaltwahrnehmung Gebrauch macht, stützt sich das Experiment auf Induktion. Ferner wird untersucht, welche dieser vier Elemente von Lorenz' theoretischem und methodologischen Ansatz sich gegenseitig begründen, z.B. ob Lorenz die Eigenschaften der Gestaltwahrnehmung verwendet, um den spezifischen methodologischen Ansatz der klassischen Ethologie zu rechtfertigen, oder ob die Methodologie der Ethologie die Gestaltwahrnehmung als wichtige kognitiven Mechanismus fordert. Es wird deutlich, dass es das Hauptziel von Lorenz' erkenntnistheoretischen Schriften der Nachkriegszeit ist, den auf qualitativen Beobachtungen beruhenden Ansatz der klassischen Ethologie zu verteidigen. Dies passt mit der Tatsache, dass Lorenz mehrmals dem Vorwurf ausgesetzt war, dass die Vorgehensweise der traditionellen Ethologie keine ernsthafte Wissenschaft sei. Selbst heutzutage ist es nicht wirklich geklärt, welche erkenntnistheoretische Rolle Beobachtung und Experiment spielen. Das Verdienst von Lorenz ist es, klarzumachen, dass wir eine Theorie der kognitiven Prozesse brauchen, die bei der Gewinnung biologischen Wissens beteiligt sind.

Theresa Schilhab Vertikales und horizontales Lernen. Einige Kennzeichen von implizitem und explizitem Lernen

Prozedurales, automatisiertes und nicht bewusstes Lernen wird immer symbolisch repräsentierten Formen des Wissens, welches verbal ausgedrückt werden kann gegenüber gestellt. Der Unterschied zwischen implizitem und explizitem Lernen besteht dabei vor allem im Vorhandensein bzw. nicht Vorhandensein von bewusster Aufmerksamkeit.

In diesem Artikel wird der Versuch unternommen unterschiedliche Lernstrategien ausgehend von einer phylogenetischen Perspektive zu untersuchen, wobei ein besonderer Schwerpunkt auf der Ablösung des Gelernten vom Kontext liegt.

Michel Hofman Von Gehirn und Geist. Eine neurobiologische Abhandlung über die Natur der Intelligenz

Als zentrales Kennzeichen von Lebewesen gilt deren "Problemlösungsfähigkeit", d.h. den Strategien welche die Herausforderungen der "Umwelt" im weitesten Sinne bewältigen. In diesem Artikel steht die sog. "biologische Intelligenz" für die Problemlösungskapazitäten von Organismen. Diese kommt vor allem in jenen Informationsverarbeitungsmechanismen zum Ausdruck, welche als Resultat der Auseinandersetzung mit den jeweiligen Umweltgegebenheiten zur Ausbildung artspezifischer "Weltbilder" führt. Darüber hinaus bedingt das Verhalten der Organismen selbst wiederum neue selektionswirksame Faktoren.

Die großen Unterschiede in den Problemlösungskapazitäten der Organismen finden ihren Niederschlag in den funktionellen Eigenschaften der neuronalen Organisation. Diese wiederum hängt von der Qualität sensorischer Mechanismen, der Informationsverarbeitungskapazität allgemein und den "inneren", zentralen Verarbeitungsprozessen ab.

Im Verlauf der Evolution kam es zu einer beträchtlichen Zunahme informationsverarbeitender Prozesse und damit auch zur Entstehung komplexerer artspezifischer Weltbilder. Damit ging die Entstehung kognitiver Eigenschaften wie Motivation, Antizipation, Emotion etc. einher. Besonders die Eigenschaften menschlicher Kognition wie Sprache und Symbolverarbeitung sind aus dieser Perspektive nicht so sehr als Resultat biologischer Selektion zu verstehen, sondern stehen im Zusammenhang mit epigenetischen Prozessen, kulturellen Einflüssen und verstärkter neuronaler Plastizität.

Steven Platek

Ein evolutionäres Modell hinsichtlich der Wirkungen väterlicher Ähnlichkeit auf väterliches Investment beim Menschen

Inn diesem Artikel werden die Auswirkungen der Ähnlichkeit zwischen Vater und Kind auf das Ausmaß väterlichen Investments untersucht. Entsprechend evolutionsbiologischer Annahmen erweist sich das väterliche Investment an den Nachkommen wegen der Unsicherheit des väterlichen Anteils an den Genen der Nachkommen durch verschiedene Faktoren bedingt.

Der in dieser Arbeit näher untersuchte Faktor ist derjenige der Ähnlichkeit zwischen Vater und Kind. Darüber hinaus werden Überlegungen hinsichtlich eindeutiger "nicht Vaterschaft" (Adoption, Stiefkinder) angestellt.