

The cradle of language: making sense of bodily connexions

Stephen J. Cowley

Introduction

Much is rotten in the ‘sciences’ of language and cognition. To those familiar with Wittgenstein’s work this is apparent in, for example, the gulf that separates investigations of mind from those of language. Equally, it appears in how empirical work tends to skate over conceptual issues while theories of discourse proceed with disregard for causal processes. Taking another direction, I invoke ‘natural history’ in asking new questions about the origins of minded and discursive behaviour. In this context, I use ‘micro-investigations’ to demonstrate how a single interaction can be used to throw light on human development. Rather than argue for the proposed ascriptions, my aim is to show the power of the method by exploring a moment when ‘understanding dawns’.

Specifically, I scrutinize an event involving a 9 month old baby. Micro-investigation serves to trace why the baby comes to *feel like* fetching a block. Having shown how neural capacities use bodily connexions with his mother’s body, I defend two claims. First, the baby develops by virtue of bio-behavioural events that index social norms. Specifically, these prompt his proto-thinking and, as a result, he has an experience of *authoring* his fetching.¹ Second, the method shows that Wittgenstein’s conceptual clarifications can be used for the real-time investigation of human biomechanics. By considering how reactive-responsive bodies use culturally-based expectations, we can generate empirical hypotheses. While concerning ‘mental’ events, these invoke –not inner minds –but bodily connexions. In cases like the one described, the challenge is to clarify how the neural control of behaviour links events to customs which are historically aligned with the ‘word’ *fetch*.

Natural history and its developmental setting

For those sympathetic to Wittgenstein, Canfield’s (1993; 1995) sketches of development are both apt and illuminating. Scrupulously avoiding cognitivism, he presents a ‘purely descriptive anthropological study’ (1993: 166) of how infants find a way into language-games. Below, emphasizing cognitive issues, I use this both in orienting the reader and sketching its limits. Specifically, I show how empirical work can generate hypotheses about how, why and what *develops*.

Focusing on what Wittgenstein calls ‘word-language’ (PI 494), Canfield treats this as “a set of customs in which words play a role” (1993: 185). Turning to the natural history of such events, he seeks out “the bedrock of the development of speech” (1993: 166). In so doing he focuses on how ‘proto-language games’ give rise to prelinguistic expression or conventional gestures that undergird our modes of life. Across time, each stage reflects “steady ways of living, regular ways of acting” (CE 397) where the conventional and verbal (‘language use’) function to meet our social needs. Speech extends “the certain action patterns that underlie its earliest uses (1995: 197)”. Indeed, Canfield thinks that, as a cultural extension of pre-existing interaction, early language requires little learning. Given the child’s dispositions, she “naturally takes part in relevant action patterns” (1995: 197) and, thus, comes to “command various convention governed ways of acting” (1993: 173). Since she can recognize

her mother, it is enough that, when together, she responds to the effects of parental alertness (1995: 198). Chimpanzees play similar proto-language games when, for example, they present a bodily area to be groomed. This constitutes a generalizable request akin to a child's special wiggle that shows she wants to get out of her chair (1993: 174). Although humans and chimps invent similar 'projects', their lives soon diverge. For Canfield, humans alone rely on stylization (1995: 202) as a basis for conventional or prelinguistic gestures (e.g. pointing). Later, drawing on animal nature, they extend these actions to verbal signs.

With verbalization, a child becomes a cultural creature. Without leaving a framework of interaction and gesture, she employs words "from a common vocabulary" (1993: 177). "One fine day", Wittgenstein (RPP II: 171) notes, she steps into language. Around their first birthday, children use words to make requests. For Canfield, it is 'brute fact' that we make, comply with, and describe wants (1993: 178). Resembling as it does the gestural expression of pre-linguistic games, there is no deep puzzle about this new word-language. The indeterminacy of translation is a non-issue because, in time, one way of acting replaces another. Looking across cultures, Canfield (1993) provides a tentative classification of language games into types: 1a) making requests; 1b) responding to requests; 2a) making intention-utterances; 2b) responding to intention utterances; 3a) uttering prohibitions; 3b) responding to prohibitions; 4) greeting; and 5) mere naming. By the end of the second year, though, cultures and groups diverge. For example, in one setting a child proceeds by saying things like 'Climbing chair' or 'Duck, frog downstairs'. Early intention-utterances begin to morph into announcements of plans and, then, branch around words like 'then' (e.g. 'Jump first, then shirt'). In later months, verbal fillings allow forward projection of action (e.g. 'eat later'), decision making ('I'll be there at 3 o'clock') and, eventually, promising ('I *will* be there at 3 o'clock').

An anthropological perspective brings much of value to conceptualizing how language influences development. By stressing that talk is fundamental in shaping how they act, Canfield shows that, early on, infants do not *learn* language. Accepting Wittgenstein's view that language is an extension of action (1976: 740), he takes the view that children –not brains –learn to participate in talk. In this way, he avoids many discussions that bedevil linguistics. His approach sidesteps debates between rationalists and empiricists as well as arguments about which aspects of word-language are 'learned' outside-in and which grow inside-out. Emphasis on how children find a way into talk makes it necessary to posit neither that brains are general learning-mechanisms nor that they run language-ready programs.² What matters, then, is how a baby comes to use utterances to act, understand and mean.

Anthropological distance makes the model simple. Above all, it distinguishes the 'natural processes' of the first months from the natural-cultural events that are made possible by conventions. With the rise of hybrid processes at the end of first year, the infant's activity meshes with word-language. So far, so good. Not only does this fit current views of human development but it provides rich description of how children use 'universal language customs'. However, given his reliance on diary records of Z's doings, Canfield is bound to emphasize speech and the speaker. Developmental effects are conflated with evolved biases and, given his method, understanding falls out of the picture. In emphasizing that children *share* adult perspectives, he underplays both the non-conventional ('Everything has a shadow except ants') and the particularity of dialogue. This happens, above all, because examining how speech can be described requires him to emphasise what is universal. The diary method also requires Canfield to play down the adult's role and, by

extension, the effects of action-around-speech. As a consequence, he says little about the child's agency or, when and why development occurs. As with all stage theories, he obscures why infants come to modify how they act. In line with Elman et al.'s (1997) critique of such models, little is said about how changes arise, why or when they occur and, especially, what it is that changes.

Given an anthropological approach, Canfield addresses neither the details of real time events nor the non-linear course of ontogenesis. By focusing on games that engage children, he can avoid questions about how children influence adults or, for that matter, when and how they come to take adult-like perspectives. By focusing on what children say, he can leave out the caregiver reinforcement and, by extension, the affective dynamics of early life. This, in turn, enables him to regard the movements and sounds of 'infant-directed speech' as a matter of *convention*. Further, it enables him to avoid dealing with those aspects of behaviour that instantiate what Wittgenstein refers to as the 'new (spontaneous, specific)' (PI, p.224).³ Similarly, while steadfast in eschewing the Inner, Canfield overlooks both what adults attribute to infants and how their behaviour draws on *practical* understanding. By treating practices as communal, tends to write as if conflict were minimal and development by 'the child' arose largely from matching habits to adult expectations. Focusing on what makes us *human*, he writes as if bodies could attune to local conventions.

While Canfield's approach helps understand how we become human, he has little to say about cognitive and causal processes. To bring such events into focus is, of course, to ask why and how development occurs. This requires one, first, to address how understanding and action mesh neural processes, affect, experience, and behaviour by others. Second, we must move beyond the observation, that while small infants or bonobos try things out, later progress is channeled by conventions. Instead we can ask about whatever-it-is that motivates the child, caregivers' prompts, and real-time interaction. In asymmetrical events each party assesses and manages the other such that the infant gradually becomes an adaptive and flexible decision-maker. Further, if language games emerge from what is spontaneous and, for the child, novel, the affect-based coupling of bodily and facial expression is the only possible basis for later action. In dealing with requests, for example, we need to investigate how, across time, children comply with, refuse and propose what adults 'require'. Equally, while projects like tearing a toilet roll into pieces (1993: 180) seem distinctly human, we will ask why. Unlike language-using bonobos, it seems, children like to draw attention to both themselves and their worlds. For example, struck by a red Santa Claus on a newly decorated Christmas tree, our son was moved to 'tell' what he could see. He pointed and, with deliberation, said 'owl'.

It is only by shifting our emphasis from what is human to individual *history* that we can examine what actually develops. Until this move is made, we can explain neither how nor why human activity changes and, for this reason, remain blind to whatever-it-is that develops. Given his focus on speech, Canfield has little to say about such questions and feels able to ignore relationships, action, content and timing. Equally, he leaves out adult beliefs, attitudes and actions or why, at specific moments, they treat children as having mental states. He fails to ask how events between people impact on understanding or how a child's later imaginings draw on the world of vocal and visual dynamics (or images). Identifying language with word 'use' implies that surface grammar (which is opaque to the child) is implicit in visible expression and vocal sound. Strikingly, little is said about how these dynamics help make the child –not just a sentient creature –but a human being who can, among other things, hope and grieve. Canfield, indeed, takes higher-order language (its verbal

features) more seriously than the sound and movement that spur early understanding. In taking a cognitive approach, I prefer to emphasise the constructive nature of intelligent action. Specifically, I use micro-investigation to show how, for one baby, sound and movement enact a request. As the infant shows, human development bases rule-following on agreements in judgment. Expanding the anthropological view, I show how the infant falls under the sway of custom. Gradually, a baby becomes a hearer-seer-actor who uses biomechanics together with culture to grasp what is wanted. In life, we say, “He *really* understands his mother”.

The cradle of thought

Within developmental psychology, most agree *when* babies begin to use ‘proto-thinking’. As Canfield implies, this arises with a cultural influence that already complements natural behaviour in the second half of the first year.⁴ In this period which Hobson (2000) calls the ‘the cradle of thought’, children use motivated pointing as well as ‘social referencing’ (Campos & Sternberg, 1981; Striano & Rochat, 2000). For example, in laboratory experiments, a visual ‘cliff’ (a special Perspex-induced perceptual effect) can elicit complex ‘triadic’ behaviour (Sorce, Emde, Campos and Klinnert, 1985). In well-replicated work, infants do not use convention to deal with the apparent danger. Rather, while younger infants balk at the visual cliff, at around 9 months, events come under the influence of social signaling. When caregivers are anxious, infants hang back but, if visibly encouraged, they may crawl forwards. The mother’s dynamic image influences how a child integrates experience with perceptions and actions. Like other child-caregiver-object activity, social referencing links two humans with an aspect of the world. While the origin of triadic behaviour is debated, its importance is endorsed by all. For Tomasello (1999), indeed, it distinguishes normal children from both their autistic counterparts and chimpanzees. To get its general flavour we can consider the events behind the following stills.

[Figure 1]

In interaction between a 9 month infant, Luke, and his mother, the first slide shows the baby inadvertently let go of a block. In the second, Sheila tries to get him to fetch it and, in the third, 9 seconds later, the infant (finally) moves off to get the block. The advantage of micro-investigation is that it can help us grasp how, why and when the baby comes to do what is wanted. Next, therefore, I put the events against a theoretical frame before using the micro-investigations to develop Canfield’s insights.

Theorizing triadic behaviour

While some invoke mental representations to explain triadic behaviour, such theories no longer dominate the field. Leaving internalist models aside, therefore, I sketch three views that, in different ways, treat word-language as a set of customs. For each model, therefore, the events contribute to the causal processes of *understanding*. The child’s relevant capacities are taken to develop because, in our species, bodies and culture are mutually adapted. However, while Shotter (2003) appeals to experience of a body-in-the world, Tomasello (1999) and Dennett (1991) take the view that culture has adapted to fit the peculiar nature of human brains.⁵

Although Shotter’s recent work aims at characterizing conscious adults, the same model can be applied to what Luke does. This is because, avoiding the Inner, he invokes two factors. First, humans are reactive-responsive beings whose ‘attitudes’

allow bodies to serve as mirrors of the soul. Emphasizing this, Shotter (1980) already takes the view that our humanity is based in 'joint action'. This is in accordance with the stills where, plainly, each responds to the other's responding. Second, drawing on Merleau-Ponty, Shotter (2003) now proposes that 'going on' derives from 'chiasmic change'. Just as vision gives us a sense of depth, Shotter claims, we experience the 'meaning' of social events. Using responsive bodies, life history makes us into meaning detectors. Applied to Luke, chiasmic change opens him to the meaning of 'fetch'. Using felt experience that is shaped by social perception, Luke grasps what Sheila's action means. Since he detects an abstract quality, he uses perceived reality and, for Shotter, fetches quite unlike a dog.

While I apply Shotter's theory to an example, he aspires to develop a general model of how shared realities arise. Many, therefore, might object at the fact that, given phenomenological grounding, it is ungrounded by neuro-behavioural events. Dennett, for example, would challenge its first-person basis by suggesting that the posited 'meaning' is a fiction based on in verbal *reports*.⁶ Given the baby's inability to understand even 'oops', Dennett would stress that it lacks adult-like phenomenology. In doing what his mother wants, therefore, the child is still a Skinnerian creature. Given lack of language, he would not be seen as undertaking anything spontaneous or new. Presumably, then, Dennett would trace the events to learning based in previous episodes of, among other things, joint attention, social play and crawling-to-fetch. On such a view, Luke's behaviour derives from general or ABC learning that, for this reason, resembles the fetching of a dog. To dig deeper, it would seem, we need models that throw light on neurobiological change.

In place of chiasmic change or ABC learning, Tomasello (1999) asserts that a special competence drives the 'nine month revolution'. Agreeing that training is insufficient, like Dennett, he emphasizes changes *within* the baby. At the same time, with Shotter, he emphasizes triadic action and, implicitly, contrasts the baby with a dog. This is the keystone of his theory: intention recognition is possible, he posits, because human brains have a 'socio-cognitive adaptation'. Literally, the child 'sees' his mother's perspective between the second and third frames. Over-riding scepticism about the causal powers of inner intentions, he might claim that this indicates infant recognition of what its mother wants (Why else would it smile *before* getting the block?). Such a view, moreover, fits with the fact that, between frames 1 and 3, the mother says 'Do you want to fetch that?' Not only does she report what she wants but, idiomatically, the child 'understands'. While using reactive-responsive bodies and ABC learning, Tomasello claims that humans have a special neural device. In spite of philosophical arguments, a baby has a private neural system that allows it to *recognize* what another person intends⁷

How do we go on?

In grasping what his mother wants, Luke does something remarkable. Although lying between the 'natural' and the 'conventional', this way of describing what happens shows nothing about why he decides to get the block. Theory tempts us to deal with this either by invoking phenomenology or appealing to a hidden process. Empirically, of course, invocation of first-person accounts is even less attractive than appeal to an inner 'competence'. This is because while appeal to hidden processes invites questions about the mechanisms of learning and evolution, phenomenology is inimical to empirical study. Accordingly, neither method can be used to deal with the *desiderata* of clarifying why, how and what develops. Equally, Wittgenstein believes, the alternatives are confused. Rather than conceptualize what is unknown as Inner, we

can examine the view that mental states arise ‘via behaviour’. How is this to be interpreted? Must we take it on trust? Alternatively, can it be treated as an empirical claim about the natural history of what we call ‘mind’ and ‘language’?

Below I use micro-investigations to scrutinize the events that constitute this moment part of Luke’s natural history. My case is that the method is powerful enough to open up new kinds of empirical enquiry. In other words, using video and audio technologies, it is possible to use observations to explore the no man’s land between the empirical and the conceptual. A person ‘observes’:

Roughly, when he puts himself in a favourable position to receive sense-impressions in order (for example) to describe what they tell him. (PI. p.187)

Using sense impressions, observational evidence can be teased apart from observational acts. Physical features of events can thus be used to develop descriptions that parallel how events are conceptualized. In scrutinizing physical acts, therefore, attention can be given to what those familiar with the relevant customs perceive in the recordings. Although sense impressions connect an observer’s natural history to his command of a word-language, the resulting descriptions are not subjective’. Rather, to the extent that observational acts draw on shared customs, they can be used in tracing how minded activity arises from behaviour. Emphasizing connexions independent of word-language, this helps bring to light how conceptualizations of the Inner distort our understanding of vocal and visible events. We come to ascribe mental capacities to others, it seems, through the ways in which behaviour is distributed in time.

Micro-investigations: the power of observation

‘Nothing is hidden’ (PI 435) can be used as a principle for asking about the rise of minded behaviour.⁸ Hypothetically, given a record of every movement that influences a child, observations might help clarify what makes us human. In taking this view, however, it must be stressed that events use –not just experiential time –but relationships, interactions and, above all, vocal and visible gesture. Accordingly, to give a sense of these complex dynamics, I develop a thick description of events by using both the video-frames shown and others from the same sequence.

Figure 2

The first two frames set the scene. As Luke sucks on the block, he finds his mother pulling at his attention. Over the long sucking time (3080 ms), Sheila distracts him by trying to set up a giving game. To this end she uses movements, gaze and accompanying vocalization.⁹ The first frame (340ms after sucking begins) shows that Luke’s gaze has already being drawn to his mother’s. By the second (2740 ms. later) Sheila has repositioned the block so that it sits on ‘top’ of her right hand. To adult eyes, the slow and deliberate movements make it ‘obvious’ that she is offering the block. Strikingly, as the second frame makes clear, Luke grasps this too: he gazes at her intensely while, broadly, matching her expression.

Luke is not simply reacting. Rather, he uses maternal movements and gaze to orient to *what she is doing*. Already, (in frame 2), he is anticipating a giving game as is shown by his hands that move towards the proffered block. In these circumstances, this has unforeseen consequences. As can be seen above the box (in frame 2), the block he was sucking falls. As it happens, the next few seconds of human interaction come to be organized around this unexpected event.

Figure 3

The 3rd and 4th frames show Sheila changes her agenda. While her gaze reacts to the falling block in micro-time (200ms), the hand movements are slower. Thus, in frame 3 (540 ms. after 2) she is still placing the block between the fingers above Luke's left hand. Her change of mind, it seems, has been triggered by her gaze-system's reaction. By frame 4 (after another 400 ms.), however, she has ceased to look at either the baby or the block she was holding. Instead, she is uttering a salient high pitched 'whoops'.¹⁰ Since this can be used to comment on errors, observers hear Sheila as taking Luke's perspective. The baby, however, is deaf and blind to this subtlety. His body, in 4, still shows readiness for the giving game. From Sheila's perspective, he 'hasn't understood'. He is disoriented in the sense that he doesn't grasp what she means. Seeking to remedy this, she slips the green block out of sight and remains in the crawling-posture that she has adopted. In orientating to the block (off camera in the bottom left), she sets a new agenda.

Figure 4

In frame 5 (700 ms later) Luke grasps what this is about. Not having responded to 'whoops' or bodily movements, a high-pitched 'ooh' acts as a precursor to Luke's gaze following. However, in looking where his mother looks, Luke has taken 9 times longer to reorient than she did. While meeting her goal, he draws on her vocalization, movement and gaze. By frame 5, then, relaxed arms show that she now aims – not to get the block – but to engage the baby in a project. In part, Luke understands, in that he gazes at the block. Although looking at the block, his arms are relaxed and there is no sign of motor movement: he does not grasp her intent. In spite of this, Luke fails to understand the block's importance (surprisingly, he has no need to look back at the block at any time over the next 2300 ms). Picking up on the point of contact, Sheila builds the object of shared attention into their subsequent activity.

Figure 5

It is hard to communicate at 9 months: for Luke seeing the block means little. Indeed frame 6 represents a moment 1280ms after 5 and a period of a further 1040 milliseconds is covered by 7, 8, and 9. Over this time, Sheila not only suggests fetching the block twice but looks from block to baby twice before, in 9, resting her gaze on the block. As before, gaze drives the interaction as she takes Luke's perspective with the words "Do you want to fetch that?" (Using 'that' to refer to what gazed at object in 6). Although he does not understand what she intends, he realizes that she wants to do something together. Thus, Luke looks back at her. Then, in 7, building on mutual gaze, Sheila varies the theme. In what we regard as a normal conversational move, his mother breaks gaze, looks at the block and says "Go fetch it". In 8, she repeats her attempt at communication but, again varying the theme, silently rests her gaze on the block. Her 'meaningful' look lasts almost as long as the others put together.¹¹

Luke's face features complex expressive dynamics. In 6, as he looks into her eyes (without understanding), lowered eyebrows hint (to her) that expects a reaction. This impression is enhanced by the fact that, in 7, he starts to close his mouth. Strikingly, when Sheila looks away, the dominant right eyebrow rises into a knitted

expression that might, in an adult, be said to express ‘what do you want’. Yet, when his mother returns to Luke’s line of sight, in 8, his gaze focuses, his eyebrows relax and, it seems, he shows the shadow of a smile. By the end of her meaningful look, in 9, she too is smiling and, like Luke, she is relaxed. The baby, though, still does not grasp her intent. Intriguingly, however, Luke’s eyebrows rise as, in 9, his smile brightens: his expression seems ‘intelligent’.

Figure 6

In frame 10 (840 ms. later), as his mother’s gaze returns, Luke is beaming. His smile has full Duchenne features that depend on a full rise of the eyebrows that, in this case, make even his tongue visible.¹² Given the slow pace of this beam (earlier, Sheila looked from block to baby and back in 340 milliseconds), this is plainly an *action*. Indeed, there is nothing public to which Luke could be reacting. After their joint effort, Luke *experiences* how to ‘go on’. Assuming that this does not come out *ex nihilo*, I describe Luke’s visible expressions of 6-9 as ‘proto-thinking’.

Figure 7

In 11 (400 ms. later), Luke sets out to pick up the block. Strikingly, moving into crawling happens twice as quickly as his smile. This hedonic event marks that the child is about to accomplish a social goal. While Luke works out how to act, this is prompted by Sheila’s movements and vocalizations. Using her dynamic image, he appropriately solves the perennial social problem of ‘What should I do now?’ Without words, he responds sensibly to ‘Do you want to go and fetch that?’ He acts within customs where, with seeming transparency, word-language makes sense. Since his smile occurs *before* fetching it gives hedonic value to ‘knowing how to go on’.

The power of proto-thinking

In aligning to “Do you want to fetch the block” Luke enacts precisely the kind of event that theories of triadic behaviour seek to explain. Next, therefore, I use their strengths and weaknesses in considering Luke’s role in ‘reinventing’ a custom. In so doing, I use a pared down view of the events:¹³

- In frames 1-4, Luke does *not* pick up on his mother’s change of project.
- In frames 5-10, he holds in mind that the goings-on concern the block (After looking at it, evocative facial expressions fade into a beaming smile).
- In frames 11-12, he grasps what his mother wants and moves off. (We lack clear evidence about what prompts the ‘decision’).

Luke’s doings cannot be explained by ABC learning because his understanding depends, overwhelmingly, on real-time events. He uses his mother’s movements which are, indisputably, dynamic attempts to get Luke to modify how he acts. In everyday terms, while using his brain to hold the brick in mind and anticipate reward, Sheila prompts him to *grasp* what she wants. The timing of Luke’s perspective-taking shows that he is only partly dependent on brain-side processes. Far from seeing meanings, recognizing intentions or relying on ABC learning, he picks up on the sense of Sheila’s behaviour. Given social experience, each reactive-responsive body prompts the other to act. Not only does this shape what accompanies proto-thinking (5-10) but, somehow, it triggers Luke’s understanding (11-12). In deciding *how to go on* the baby shows exquisite sensitivity to affectively charged action. Given powers of gaze following, inhibiting alternative actions, and keeping to the topic, this need be

complemented by little more than well-timed felt action. Whether or not the biases have a genetic basis, some (perhaps, all) internal change fits Dennettian logic. Using ABC learning, Luke meshes what comes automatically with behaviour that, in the circumstances, enacts what his mother wants.

In a case like this public events are of paramount importance. What Luke does, therefore, is incompatible with an inner-process view of understanding. The incident thus jars with both Tomasello's inner competence and the chiasmic change posited in Shotter's appeal to the phenomenal 'I'. Events like these depend on, not just Luke's brain, but how joint action is *distributed in time*. The fact that Luke takes 9 seconds to grasp what his mother wants shows that, biomechanically, grasping how to go on is not a simple task. Indeed, the pace of events itself speaks powerfully against any model based in phenomenology or intention-recognition. Second, because of the slow pace, we can trace he draws on neural abilities for not forgetting, following gaze, inhibiting action and so on. In spite of his brain work, however, Luke fetches under *social* influence. Unlike a dog, he uses his mother's dynamics and, by extension, needs no fetching inclination (see 5-10), no orienting excitement, and no canine quickness. Unlike a Skinnerian creature, Luke uses Sheila's *belief* (viz. that Luke may be able to fetch the block). While training, reinforcement, and imitation contribute to the experience, Luke is sensitive to how it is enacted. Fine control enables him to integrate events as disparate as Sheila's expression, the brick, and feedback from his smile. Far from needing to see meanings or recognize intentions, brain-based learning allows events to be concurrently controlled by both parties. Just as a small baby can use its mother's manifest wishes in learning when to fall silent (see, Cowley et al. 2004), *fetching* arises as two brains exert dual control.

Luke's beaming smile goes some way towards explaining what we might call 'mind-reading'. This arise because, given slow emergence, it prompts and rewards both parties. For Sheila, smiling in advance shows *real* understanding and, for the baby, it provides a sense of authorship.¹⁴ Indeed, the smile both *precedes* getting the block and rewards Luke's previous (in)activity. Leaving much aside, the timing of the smile shows that Luke's 'action' results from integrating proto-thinking with maternal movements. Second, interindividual coupling prompts Luke to *feel* when to exploit how Sheila enacts a want. Development thus co-occurs in three domains. It changes the baby's brain and body, shapes what is likely in a relationship and, crucially, alters Sheila's beliefs. Indeed, this is what spurs her to act in ways that make proto-thinking a prelude to a reward. Using world-side events, the baby smiles and, at once, enacts his understanding. Whatever the neural basis of the events in 11-12, Luke *wants* to get the block. For the baby, the feeling is spontaneous.

Since much occurs world-side, spontaneity must be distinguished from novelty. First, all Luke's actions are singular in spite of the fact that he relies on practiced routines. The novelty of his action, then, arises from connecting with his mother's doings. It takes on a value as, in real-time, he aligns to what she says and does. In contrast to spontaneity, novelty is thus rooted in culture. By linking these with what the baby does, *Sheila* makes it seem that Luke 'really' understands. If, she accepts this, it becomes reasonable to test it (and train Luke) by setting up new 'fetching' opportunities.¹⁵

While spontaneity is jointly constructed, it is experienced individually. The hedonic property of the act is experienced by the baby (through the beaming smile) even though, as micro-observation shows, its causal basis lies partly in a set of customs. Without the cultural context, Luke would be unable to align his behaviour to what is actually said. He exploits not just dynamic attunement but, above all, how

Sheila acts in getting Luke to fetch the block. The social event thus depends on close-coupling between intentional movements and the baby's rudimentary ability to exploit caregiver doings. Remarkably, this gives shape to an event whose novelty and spontaneity put Luke on the edge of language. Bio-behaviourally, changes in his repertoire are based in how he times his part in joint action. While *fetching* is easy for a crawling, grasping baby, the event's coherence depends on Sheila. Remarkably, she enables the baby to align its behaviour with the grammar of 'fetch'. Using culture, Sheila helps Luke attune to a bundle of local customs.

Since development integrates world-side and neural events, the process can be seen as based in mutual *gearing*. Using both micro-movements and longer lasting 'actions' each party adapts to the doings of the other. Thanks to this gearing, future interactions come under the influence of both true and false beliefs. In this case, Sheila's perception of Luke's pleasure makes her likely to encourage fetching games because, as she might say, 'he understands what to do'. As Luke becomes skilled in grasping where 'fetch' serves his interests he learns about 'requests'. Mutual gearing can produce such outcomes because it is based on bodily action where adults use cultural constraints (expressed by the words actually spoken). Luke can adapt to these by using both extant skills and his sense of authorship. Further, in that causal chains reach beyond the skull, Luke and Sheila develop ways of playing based on joint control. In short, natural history can draw on bio-behavioural events in a cultural setting. The example clarifies how requests arise from brute facts (Canfield, 1993; 179). Given cultural constraints, decisions about action use brain-side processes that prompt bodies to gear to each other while setting off neural events that drive joint activity and, at the same time, affective experience.

Connexions before language

Many concur with Bennett and Hacker (2003) that scientific enquiry cannot be used to investigate mental phenomena. In their view this is because, as Wittgenstein shows, it is mistaken to regard the relevant predicates as naming properties of minds (or brains). Thanks to a Cartesian and empiricist heritage, such misconceptions dominate everyday psychology and, if used in science can, they argue, only compound confusion. This happens, above all, if we conceptualise the mental as private, as available to introspection and as constituted by inner systems that correspond to named mental entities. For Bennett and Hacker, such views constitute an important 'misconception of the logical nature of experience and its ascription' (2003: 85). To avoid the error, we must distinguish two forms of enquiry: either we must investigate logical relations or, if we wish to pursue empirical work, we should investigate the nervous system. If science seeks to clarify our mental concepts, it seems, we are bound to fall into conceptual confusion.

Must the scientist leave the mental to philosophy? Does minded behaviour falls outside the domain of empirical enquiry? While some theorists may take this view, Canfield opposes any such interpretation. On the contrary, he thinks that reducing Wittgenstein's work to arguments about concepts can produce superficial readings. To establish if Wittgenstein is correct, he thinks, 'what one must do is look below the surface' (1986; p. 153). Presumably, this is why he asks how children come to participate in customs where words play a part. Breaking with Wittgenstein's precedent of 'inventing' natural history, his anthropology draws on empirical evidence. Further. As argued above, he *establishes new facts bearing on the social basis of mind and language*. The spirit, moreover, is consistent with much in Wittgenstein. For example, he is forthright in suggesting that his analysis does not preclude investigation of 'the possible causes of the formation of concepts'.¹⁶This same goal, of

course, drives both Canfield's work and my micro-investigation of how Luke develops practical understanding of *fetch*. Conceptually, moreover, these approaches remain, in some sense, Wittgensteinian. This is because, in contrast to Dennett and Tomasello, they reject pure third person accounts while, unlike Shotter, making no appeal to phenomenology. Micro-investigation of Luke's fetching brings our facts that echo both these perspectives and, crucially, his mother's second person view. Luke gains practical experience of *fetching* because he engages with Sheila's beliefs and what she ascribes to him. This is only possible, of course, because of how his body (and brain) respond to events, giving him –among other things- a feeling of what happens. Folk views of 'understanding' cannot therefore help us think about how such a concept develops. This indeed is why I emphasise that learning to fetch –and learning to talk –are grounded in mutual gearing. Much mental life arises as babies show sensitivity to behaviour while using neural tricks that allow them, for example, to follow gaze, hold an object in mind, and feel like the authors of their actions.

Empirically, micro-investigation sustains the gearing *hypothesis*. This is, for current, purposes sufficient: in this paper, my aim is to show how empirical work can be used to develop new perspectives on the rise of mind and language. As with other independently observable phenomena, simple concepts can be grounded in behaviour. What Luke shows, from this perspective, is why we must reject the individualism of both mainstream cognitive psychology as well as conceptual constructs that give coherence to mental talk. Even if this leads us to see human subjects as bodies who, immersed in culture, develop a range of potentialities (2003: 52), I doubt that Bennett and Hacker would approve of the conceptual implications of this work. This is because, if some concepts arise as bodies gear to each other, this contradicts our everyday views of 'understanding'. While consistent with the claim that neither minds nor brains function causally, learning to fetch –and other practical knowledge – requires special brain mechanisms. It implies that social events depend on neural processes that exploit intricate processes of dual control.

Minded behaviour emerges because of how changing brains exploit events in a micro-world. Mind and language exploit connexions between bodies that are, traditionally, overlooked in both commonsense and cognitive science. In natural history, asymmetries between baby and caregiver may play important developmental roles. Indeed, it is difficult to see how a baby could gear to conceptually-based expectations without special predispositions. Further while a baby learns conventions, these are of less intrinsic interest than the modicum of control these provide. Through mutual gearing, using biomechanics, Luke becomes a player of intentional games. This, moreover, is entirely compatible with the developmentalist's finding that qualitative change co-occurs with the emergence of triadic activity. What micro-investigations add is a method that, in this case, serves to generate hypotheses about both proto-thinking and how babies use two brains to develop new forms of control. Although a little *person* experiences 'how to go on', Luke's feeling arises as events in his brain enable him to co-ordinate with his mother's culturally constrained dynamics. He uses how her 'want' shapes her looking, moving and vocalizing to create opportunities for action that will impact on his own development. Given such findings, one must contest models that define development around the organism. In the rise of triadic behaviour, a baby's sensitivity to bodies precludes any need for either a socio-cognitive device or a first-person phenomenology.

Fetching links biomechanics, maternal beliefs and a set of customs. It is part of a developmental process that depends on how a child's body can use caregiver activity to, among other things, produce a sense of agency.. Indeed, Luke's doings

seem to be highly consistent with what Wenger (2003) calls the ‘illusion of conscious will.’ Grasping when to get a block is a form of affective experience that makes Luke feel like a subject and, where similar events recur, this will help him to anticipate what is the right thing to do. Far from seeking to pursue such empirical hypotheses, however, my goal is to show that, drawing on Wittgenstein’s conceptual work, we can examine natural history by micro-investigations. In some cases at least, minded behaviour arises as maternal beliefs prompt a baby to discover how sense-making can be prompted by bodily connections.

A set of customs

Everything about Luke fits the view that language is a set of customs in which words play a part. Early on, word-language can only function by virtue of how it is embedded in behaviour. By implication, models where learning to talk is conceptualized in terms of learning or acquiring word-forms are misguided. Rather much early development proceeds as bodies act to establish agreement in judgments. In the example, Luke and his mother need merely concur, in practice, that *this* affords fetching a block. That is all. The baby’s ability to *go on* integrates biomechanics, proto-thinking and beliefs based in a set of customs. By mutual gearing, Luke and his mother each prompt *de facto* judgments. Rather than appeal to stylization, speech and the individual speaker the example gives a sketch of how, given culture, mind and language emerge from biomechanically mediated events.

Once Sheila believes that her baby understands *fetching*, she will encourage such activity. Indeed, this strange capacity can shape much of our natural history. Exotic beliefs can be used to train infants by getting them to gear to our manifest expectations. This is brought out in how Sheila persists in getting Luke to fetch the block until, after a full 9 seconds, the baby meets her expectation. Mutual gearing not only shapes their joint action but, remarkably, enables Luke’s brain to mark the event as future relevant. Far from needing word-language, affect enables him to experience *fetching* as spontaneous, novel, appropriate and rewarding. Similar gearing will, at other times, give rise to other joint practices. Without needing conceptual knowledge, a baby can rely on how other persons exploit word-language. Far from using Inner understanding, biomechanical and affective tricks enable Luke to align his actions in events that can elicit detailed word-based descriptions (e.g. fetch, block). His decision making uses Sheila’s manifest thinking: his categories are grounded in *actions based on a history of dual control*. Given his mother’s view of their agreements in judgment, mutual gearing can lead to the (re)-invention of intentional games. As Luke becomes a person, he will hone many skills. Just as gearing can promote fetching, other settings can promote capacities useful in practices as diverse as word-language, fighting, singing and dancing.

1 As explained below, the phrase applies to some of the baby’s expressive moves in a fetching game. My aim is not to delve into how the biomechanical connects with the social but, echoing Wittgenstein, of ‘noting’ what happens (PI 654-655). Of course, appeal to proto thinking parallels Canfield’s (1993:173) invocation of ‘proto language’.

2 Pinker (1994) and Sampson (1996) debate rationalism and empiricism (for another view, see Cowley (1997)). Hirsh-Pasek and Golinkoff (2000) discuss language acquisition by contrasting outside-in and inside-out models.

3 Canfield uses this quotation (1995: 2003) to consider the rise of complex language-games. In contrast, I focus on how a child can author behaviour that is *felt* to have these properties.

4 Elsewhere I use micro-investigation to show signs of culture in a child of three months (see Cowley *et al.* 2004).

5 While Dennett writes little about development, like Tomasello he plays down the power of human bodies. Shotter, by contrast, ignores the brain. Each echoes the brain-body dualism of much work on cognition (Bennett and Hacker. 2003).

- 6 Irrespective of views on Dennett's heterophenomenological method, the claim that first person reports are, in part, fictional fits with biophysics. Thus, in work on how we see a sequence of images as a moving film, Kolers and von Grunau (1976) report experiments on the phi phenomenon. Subjects consistently report seeing a red light that, at a determinate place, becomes green. Physically, however, a red light flashes (completely) before a green one so. The impression and reports arise because the flashes occur only few milliseconds apart. For Dennett, therefore, even simple perceptual first person 'reports' are intrinsically fictional (see Cowley and Love, in press).
- 7 While never invoking Augustine's 'natural language of mankind', this too requires a brain that shows us the intentions of others. For critique of Tomasello's theory, see Cowley (2004).
- 8 Wittgenstein uses 'nothing is hidden' to suggest that how sentences represent meaning is open to view. Using micro-investigations, I apply the same insight to action (cf. PI 126). Even if a full bio-behavioural theory would depend on using standard empirical methods, micro-investigations are needed to counter folk beliefs and, by so doing, ground plausible hypotheses.
- 9 She seems to say, softly, 'go on my baby love'.
- 10 At this level, people cease to be unitary: gaze, hand movement, and vocalization seem partly autonomous.
- 11 The durations are 340, 720 and 1040 milliseconds respectively.
- 12 In 1862, Duchenne pointed out that while half-hearted smiles involved only muscles of the mouth, the sweet emotions of the soul also activate the pars lateralis muscles around the eyes. For a recent neurally informed view, see Soussignan (2004).
- 13 This is necessary in events that ground agreement in judgments. Since these warrant many ascriptions, it is easy to get bogged down. I hope readers will agree that the 'pared down' account is broadly correct.
- 14 While the smile might show that Luke recognizes an intention (or sees a meaning), this is doubtful. First, recognition is passive and – typically – does not produce smiling. Second, Luke gives a full Duchenne smile. Third, this ascribes causal power to inner intentions. Fourth, since getting blocks (unlike, say, bones), is not intrinsically pleasurable, the smile is more likely to be consequence of coming up with a way to act.
- 15 Although this is probably not the first instance of such fetching, Luke is not yet skilled in the activity. For expository purposes, therefore, I regard it legitimate to write as if this were a 'first'.
- 17 For *philosophical* purposes, of course, Wittgenstein rejects enquiry into the 'possible causes of the formation of concepts'; instead, he 'invents' accounts of natural history (PI p.230). One implication of this paper is that mothers too invent natural histories for their babies just as, later, children narrate their life-histories.

References

- Bennett, M. & Hacker, P. (2003) *Philosophical Foundations of Neuroscience*. Oxford: Blackwell.
- Campos, J. & Sternberg, C. (1981) Perception, appraisal, and emotion: the onset of social referencing. In M. Lamb & L. Sherrod (eds.), *Infant Social Cognition: Empirical and Theoretical Considerations*, pp. 273-314. Hillsdale, NJ: Lawrence Erlbaum.
- Canfield, J. (1986) 'I know that I am in pain' is senseless. In J. Canfield (Ed.) *The Philosophy of Wittgenstein, Vol 8: Knowing, naming, certainty and idealism*. Garland: New York, pp. 113-129. {Originally published, 1975}.
- Canfield, J. (1993) The living language: Wittgenstein and the empirical study of communication. *Language Sciences*, 15/3: 165-193.
- Canfield, J. (1995) The rudiments of language. *Language & Communication*, 15/3: 195-211.
- Cowley, S.J., (1997) Conversation, co-ordination and vertebrate communication. *Semiotica*, 115 1/, 27-52.
- Cowley, S.J., Moodley, S. and Fiori-Cowley, A. (2004) Grounding signs of culture: primary intersubjectivity in social semiosis. *Mind, Culture and Activity*, 11/2: 109-132.
- Cowley, S.J. (2004) Simulating others: the basis of human cognition? *Language Sciences*, 26/3: 273-299.
- Cowley, S.J. & Love, N. (in press). Language and cognition or how to avoid the conduit metaphor. To appear in U. Oskulska & A. Duszek (eds.) *Bridges and Walls in Metalinguistic Discourse*. Berlin: Peter Lang.
- Dennett, D. (1991) *Consciousness Explained*. Boston, MA: Little, Brown.
- Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996) *Rethinking Innateness: A Connectionist Perspective on Development*. Cambridge, MA: MIT Press,
- Hirsh-Pasek, K. & Golinkoff, R. (2000) *The Origins of Grammar: Evidence from Early Language Comprehension*. Cambridge MA: MIT Press.
- Hobson, P. (2000) *The Cradle of Thought*. Oxford: Oxford University Press.

-
- Kolers, P. & von Grunau, M. (1976) Shape and color in apparent motion, *Vision Research*, 16: 329-335.
- Pinker, S. (1994) *The Language Instinct: The New Science of Language and Mind*. London: Penguin.
- Sampson, G. (1997) *Educating Eve*. London: Cassell.
- Shotter, J. (1980) Action, joint action and intentionality. In M. Brenner (ed.) *The Structure of Action*, pp.28-65. London: Blackwell.
- Shotter, J. (2003) Cartesian change, chiasmic change – the power of living expression. *Janus Head: Journal of Interdisciplinary Studies in Literature, Continental Philosophy, Phenomenological Psychology and the Arts*, 6(1), pp.6-29.
- Sorce, J, Emde, R.N, Campos, J.J. & Klinnert, M.D. (1985) Maternal emotional signaling : Its effect on the visual cliff behavior of one-year-olds. *Developmental Psychology*, 21: 185-200.
- Striano, T. & Rochat, P. (2000) Emergence of selective social referencing in infancy. *Infancy*, 1/2: 253-264.
- Tomasello, M. (1999) *The Cultural Origins of Human Cognition*. Cambridge MA: Harvard University Press.
- Wegner, D. (2003) The mind's best trick: how we experience conscious will. *Trends in Cognitive Science*, 7: 65-69.