The lazy mindreader : a humanities perspective on mindreading and multiple-order intentionality

Duijn, M.J. van

Conclusion

In the Introduction I have put forward a series of observations and questions, and set out as the overall aim for this thesis to work these out in detail, thereby rethinking the nature of the complexity posed by intentional states underlying human interaction events. Throughout the subsequent chapters, I have looked at how such complexity should be thought of *conceptually* and how it can be handled *linguistically*. Also, I have made suggestions about how human agents are able to process it *cognitively*. I have pointed out that this thesis' methods and frameworks are ultimately rooted in the humanities, but that its topic requires analysing and bringing together insights and material from areas across both the humanities and sciences, in particular linguistics, literary studies, philosophy, and various disciplines within the social, cognitive and biological sciences. Therefore I have characterised the overall project as being an example of *topic-oriented scholarship*: instead of working with the questions, assumptions, and methods common in one academic discipline, I have applied a pragmatic mix of expertise and methodology I considered suitable for making progress on the chosen topic.

The observations and questions put forward in the Introduction are grouped in thematic rubrics: complexity posed by multiple intentional states, representation of complex thoughtscapes in discourse, communicative and cognitive economy, and the consequences for existing theories and frameworks, in particular our evolutionary story. Given that Chapters 2-6 already end with a section summing up discussion points and conclusions, I will structure what follows using these rubrics rather than going through the final remarks in the "chronological" order of the chapters. However, before doing so I will first provide an overview that will pull together the lines connecting this thesis' main findings and insights.

Overview

Human interaction is characterised by an endless "polyphony", a perpetuated multitude of perspectives and perspectives-on-perspectives. Not only do we share and coordinate our own inner life with that of the people we interact with, but we also constantly make implicit and explicit reference to the intentional states of others who may or may not be present at the time of speaking, or who may even exist only in the imagined worlds of thought and fiction.

In philosophy of mind and the cognitive sciences, this polyphony of perspectives has often been conceptualised as a series of *embedded layers* of the form "A *thinks* that B *understands* that C *expects*...etc.". Building on this tradition, tests have been devised targeting what is referred to as "multiple-order intentionality" or "mentalising", generally finding that humans can handle up to around five layers of embedded complexity. This idea was subsequently implemented in theories and frameworks concerning such topics as cooperation, moral reasoning, social aptitude and variations in quality and size of individual social networks, (a)typical development, language competence, and appreciation and production of literature.

However, throughout this thesis I have suggested that the conceptualisation of mentalising involving series of embedded layers stands in stark contrast to how dealing with a polyphony of intentional states takes shape in actual discourse and interaction. Firstly, it appears that intentional states are normally connected and interlinked in all kinds of different ways, forming what I have termed a "thoughtscape" rather than a recursive string. Secondly, if a (complex) thoughtscape is being represented in natural discourse, the labour of indicating the connections between intentional states is generally distributed over a wide variety of linguistic elements across lexical, grammatical, and narratological categories. Hence, representing a thoughtscape by a proposition featuring only recursive embedding of clauses frequently yields an infelicitous rendering of the actual relationships between the intentional states. Thirdly, parts of the thoughtscapes underlying interaction events come "packaged", and they are unpacked only if the context so requires. Fourthly, instead of thinking of human interaction as a process in which individuals seek to "join" their

intentional states through working out what others *intend* one to *understand* that they *want* one to *believe*, and what one *intends* others to *think* that one *intends* them to *believe*, such interaction should rather be seen as a joint activity in which cues are provided negotiating how a set of already shared beliefs (referred to as the common ground following Clark, 1996) should be updated. Fifthly, linguistic items, ranging from lexical units to grammatical patterns and narratological structures, embody a wealth of experience accumulated from generations of language users attempting to coordinate their own intentional states with those of interlocutors and third parties, regarding the non-intentional world by which they are surrounded. As such, linguistic items can be thought of as "supercues" supporting the process of negotiating how a thoughtscape should be conceptualised: depending on context and signaller's goals, language can offer precise analytical tools for working out details and nuances of how intentional states are mutually connected, or, conversely, language can offer cues referring to entire chunks of a thoughtscape holistically, leaving such details and nuances packaged in order to serve convenience and efficiency. In this way, the "toolkit" offered by the symbols of a language contributes at once to the richness and detail of the total system, as well as to its economy and workability. Sixthly and finally, by acquiring the lexical, grammatical, and narratological conventions that constitute this toolkit children become full-blooded human mindreaders—who are as a rule *lazy* mindreaders: operating in a socio-cultural environment that contains the coagulated interactional experience of earlier generations, saves processing costs on the individual level.

Complexity posed by multiple intentional states

As an alternative for conceiving of intentional-state complexity as a series of embedded layers, I have suggested the notion of the *thoughtscape*, defined in Chapter 1 as the total network of interlinked intentional states that are in some way relevant in the course of an interaction event. With respect to this, there is a question brought up in the Introduction that has not yet been answered: can the concept of the thoughtscape function as an alternative measure of intentional-state complexity? In this section I will suggest that this is possible,

but that an accurate implementation of the thoughtscape as a measuring tool requires further investigation of the qualitative differences between different sorts of intentional relationships.

The idea of embedded layers clearly comes with a natural complexity scale: counting layers and comparing numbers of layers across tasks, individuals, species, and so on, seems to be inherent in the very concept of multiple-order intentionality. However, I have pointed out that this is problematic: when the orders of intentionality are used as a meta-linguistic measuring tool to form paraphrases of how intentional states underlying an interaction event are related, they often force these intentional states into a strait-jacket that misrepresents the actual complexity. Paraphrases featuring only embedded clauses are at the same time too complex *and* too simple: on the one hand, they easily produce opaque renderings of a situation that is not too hard to grasp as such or in the form of a story, while, on the other hand, such propositions often fail to convey all kinds of nuances and details. A good example is Zunshine's paraphrase discussed in Chapter 3, which is indeed much harder to understand than the relevant passage of the novel itself is, while at the same time underrepresenting the subtle perspective shifts characteristic of the 360-degree view offered by Woolf's prose. Even in the case of *Othello* this holds true: as I have argued in Chapters 2 and 5, this play has the quite exceptional feature of meaningfully embedding a relatively large number of perspective layers. However, this only works because its plot combines multiple scenarios that come with significant *differences* in shared knowledge or common ground (scheming plans, revenge, adultery), and, as argued in Chapter 5, it is precisely in those cases that it is relevant to work out what A *knows* that B *intends* that C *thinks*...(etc.) In that sense, the embedded proposition "Iago *intends* that Cassio *believes* that Desdemona *intends* that Othello *considers* that Cassio *did not intend*..." *does* in some way accurately summarise the situation with which the audience is confronted after the first two Acts. Even so, this proposition suggests a sense of opaque complexity that is unlikely to be perceived when following the play's narrative on the one hand, while being far too narrow to catch the full complexity of the thoughtscape that has emerged by the end of Act II on the other hand.

In the case of both *Mrs Dalloway* and *Othello* intentional-state complexity was better conceptualised as a thoughtscape than using a proposition featuring a string of embeddings. However, although the relevant parts of both texts were misrepresented by such a proposition (that is, the first more so than the latter), it offered an easy way to *compare* the two in terms of the assumed intentional-state complexity, simply by counting the number of embeddings. Such counting is much less straight-forward when comparing two thoughtscapes. To recapitulate, in Chapter 1 and 5 I have discussed that the basic unit of a thoughtscape is the interaction event, in which two or more parties interact using linguistic and/or non-verbal cues. Prototypically, these are interlocutors in a face-to-face setting, but in the special case of fiction the position of the speaker is taken by the narrator and that of the addressee by the reader, hearer, or spectator. Speaker and addressee have a common ground, a set of shared beliefs, which they update in the course of the interaction event. This does not by default include reasoning about intentional states, but both the intentional states of the speaker and addressee and those of third parties can become relevant depending on setting and context (whether they are subsequently spelled out in language or left implicit is another question—see the next section). For example, when a narrator tells a story about two characters, it makes sense to use linguistic items that invite the reader to form representations of these characters' intentional states at various points in the development of the plot. These intentional states are both embedded in the perspective[102] of the narrator, but not necessarily into one another. They can exhibit causal relations (A *thought* X, which made B *want* Y), form meaningful conjunctions from the perspective of the reader (A *thought* X and B *thought* Y), or, indeed, be embedded (A *believes* that B *thinks* that X). A thoughtscape is therefore, as I have argued, best conceptualised as a *network* of mutually linked intentional states. This network emerges in the course of an interaction event to the extent that particular contexts require interlocutors to draw inferences about each others' or third parties' intentional states. For example, the context

[102] At the beginning of Chapter 4 I have made the distinction between "intentional state/mindstate" and "perspective/viewpoint" roughly as follows: intentional states or mindstates are atomic "snapshots" of a subject's relation to an object; a perspective or viewpoint comprises the broader total of an actor's subjectivity of which intentional states are isolated parts or "snapshots".

of a practical joke may require A to make B believe that X is the case, while at the same time telling C that B thinks Y; or, in the context of a misunderstanding A may wonder what B intended A to think when saying X.

With this in mind we can turn to the question of measuring intentional-state complexity. Having the form of a network, thoughtscapes can in principle be used for this purpose: counting the number of intentional agents and relationships between them (i.e. the number of edges in a network graph) can work as a global *quantitative* measure of thoughtscape complexity. However, the practical implementation of this measure is not straightforward, given that one would also have to account for the fact that the relationships (i.e. edges) can be *qualitatively* different in complexity: "A *thinks* that X and B *thinks* that Y" may be easier to process than "A *suspects* that X and B *does not think* that Y" or "A *wants* B to *believe* that X". Getting a grasp on such qualitative differences can begin by incorporating existing psycholinguistic insights into processing grammatical procedures such as negation or anaphoric reference into the framework of the thoughtscape. In this way, single relationships can be compared and weighted for the amount of processing effort they require. On top of that, it will be necessary to develop new research that evaluates the processing implications of incrementally integrating multiple intentional relationships. With the "old" conceptualisation of complexity as embedded layers it seemed evident that every added layer led to an increase in cognitive load, until a limit was reached at around five orders (or, according to some authors, a higher number—see Chapter 6 for a discussion). Contrastingly, it is not clear whether every additional edge in a network of intentional states should be seen as a factor adding cognitive load, and whether this adds up to a maximum network size in terms of what individuals can on average process cognitively. In Chapter 2 and 3 I have suggested that it is not plainly the case that cognitive factors limit story complexity, but rather that understanding reaches as far as the maximum complexity that can be covered by a story. At a glance this may seem a sophistic twist instead of a real claim, but think of it this way: if the reader's, hearer's, or spectator's understanding of a network of intentional relationships depends for an important part on factors inherent in the language and narrative used, it must be assumed that stories employing a more optimised mix of such factors can represent more complex networks. In this way, the limits of what is the *most*

complex network that can be covered by a story can be pushed until, at some point, the story "breaks down" in terms of coherence, becomes unreadable, or in another way fails to do the job of exposing its plot to its addressees. In this view, the upper boundary to handling intentional state complexity is not *given* by individual cognitive limits that have an average height in human populations, but instead dynamically *produced* by an interplay between individual cognitive factors and group-level cultural phenomena of language and narrative. Such a view has consequences for the evolutionary story set out in Chapter 1 and fits with two other pieces of the puzzle: one pertaining to linguistic symbols as "supercues" and the other to cognitive economy.

Representation of thoughtscapes in discourse

Far from all intentional states relevant to interaction events are highlighted or made explicit. Consider once again the picture with the woman on a staircase carrying a suitcase in Figure 1 of Chapter 1 (it can be debated whether standing in the position of the photographer and deciding to offer a helping hand counts as "interaction", but let us for the sake of the argument assume that it does). Within this interaction event as seen from the perspective of the person standing in the position of the photographer (referred to as the *mindreader* in Chapter 1), there is a role for the intentional state of the woman with the suitcase: an estimation of this intentional state (the *mindread*) is one of the factors on the basis of which the mindreader decides whether to take action.[103] It is possible to formulate an explicit paraphrase of the mindread (e.g. "mindreader *thinks* that mindreadee *intends* to carry the suitcase upstairs"), but in most cases the mindread will remain an implicit factor in a decision about future behaviour, which can be non-linguistic (provide a helping hand) or linguistic (e.g. "Can I help you?"). Only in exceptional cases will an utterance in the practice of everyday interaction come close to an explicit paraphrase. In this

[103] Later on in Chapter 1 I have nuanced this view by suggesting that this can be safely said on the W-level of what the task is, but that caution is needed on the H-level of how the task is carried out: in this context, stating that the intentional state of the person on the staircase is a factor on the W-level does not entail that the process "running" in the mindreader's mind necessarily resembles that in other contexts in which mindstates play a role on the W-level (see Section 1.1.4).

example that could happen when the mindreader appears to have made the wrong estimation after all: imagine that the mindreadee wanted to test whether the suitcase would be too heavy to be carried around for a week. In that context, a conversation could be imagined in which she declined help, after which the mindreader said "I *thought* you *intended* to carry it upstairs…"

I have argued that interaction events should be seen as a negotiation process between speaker and addressee of how a (presumed) set of shared beliefs or common ground should be updated. Whereas it is in principle possible to carry out such negotiation without language, language comes with powerful, specialised cues (or *tools*) on several levels (lexicon, grammar, narrative) to facilitate this process: from packages that project entire scenarios at once (e.g. "Sorry, I *misunderstood*") to all kinds of tools for indicating subtle perspective shifts, epistemic stances, and so on, to work out the details (e.g. "I *see*! You *wanted to try* whether…but to me it *looked* as if…I just *wanted* to be helpful…"). It is significant to notice that a context in which the help is appreciated goes much more naturally without words than one in which help is declined. In the first case, both parties are "on the same page", whereas in the latter case there is clearly a difference in how they envisage the interaction to continue—this once again suggests that contexts with *differences* in common ground are the ones in which working out who-thinks-what is relevant.

In Chapter 5 I have argued that the dynamics of negotiations about how to update the common ground can be pictured as having three dimensions, with cues serving coordination between the intentional states of signaller and addressee ((*x*)-axis), with respect to objects of joint attention ((*y*)-axis), and with respect to third-party perspectives ((*z*)-axis). From a "synchronic" perspective looking at one single interaction event, each linguistic item supports this process because both interlocutors share knowledge about how it operates on one or more of these axes (i.e. their *meaning* in this model). From a "diachronic" perspective, however, it can be said that every lexical item and grammatical procedure ultimately is the result of generations of language users trying to coordinate their mindstates in interaction with each other and the environment, thereby converging on solutions that are communicatively effective, cognitively efficient, and learnable for new generations of language users. In this sense, when we use language today, we have "supercues" at our

disposal that embody the accumulated interaction experience of generations of language users.

Evolution of the lazy mindreader

The advocated view allows individuals to be "lazy" in multiple ways. In Chapter 5 I have argued that Scott-Phillips (2015) and Sperber (2000) misconstrue the complexity of the mindreading abilities needed for present-day human interaction *in theory*: rather than working at several levels of intentionality by default, individuals make the a priori assumption that a set of knowledge states (the common ground) is mutually shared. In terms of investing processing effort, this allows them to start from zero and scale up if necessary.[104] On top of that, I have suggested various mechanisms and principles that *in practice* save individuals from having to scale up most of the time, including the observation from relevance theory and the study of alignment that interlocutors operate "in complementary predicaments" (Apperly, 2011: 115, referring to Sperber and Wilson, 2002), and the idea that most everyday interaction exhibits a structure of testing-adjusting-retesting in which representations have to be just "good enough" for the interaction to continue, but no better. These views were also consistent with observations made in Chapter 2 and 3, where the focus was on the exceptional situation posed by some works of (literary) fiction: rather than being challenged to the extreme by the necessity to keep track of complicated networks of intentional states in order to follow the plots, it appeared that the reader could "sit back and relax" while the narrator brought a rich mix of strategies to bear gradually exposing all the involved intentional states and their mutual relations.

A different part of the story is that human children can not only use linguistic elements they acquire for communication, but they presumably also extend their thinking repertoire through the process of learning how to use

[104] Note that this view also solves the "infinity paradox" mentioned briefly in the Introduction. Interlocutors normally do not have to "meet" the other's mind a few mindreading steps down the line (in which case there are always more steps possible—the infinite regress, as many authors suggest); instead, they are already "together" at zero by default. If each step is jointly taken from there (through communicative negotiation), there is no longer a reason why an infinite regress would be lurking.

some of them. It is important to note that this works for all individuals within a particular cultural-linguistic community *in the same way*: given that A and B are both (adult-level) members of the same community, individual A can not only presume a particular set of beliefs to be mutually shared with B, but she can also be sure that B has the same "toolkit" on board for negotiating how these beliefs should be updated or how differences in common ground should be detailed. Examples of tools for such negotiation (as discussed in Chapter 4 and mentioned in the former section) include viewpoint packages, supporting the possibility to project entire scenarios holistically, and various grammatical patterns making it possible to work out the details of who-knows-what one by one. If A uses the word "mislead" she at the same time activates in herself and in B the complete topology packaged by this word, including the roles of a misleading and a misled party and the default intentional relationships between them. Both interlocutors can benefit from this in the next steps of their interaction by integrating the topology with context-specific details.

This view also implicates a special niche for narratives of all kinds, ranging from the day's latest gossip to myths, parables, and even literature. Acquiring the tools for negotiating and coordinating intentional stances comes down to learning how they are used in different contexts. Stories of all kinds provide a rich and varied environment for such learning. Whereas "mislead" may perhaps occur frequently enough in everyday contexts for language learners to acquire its full topology, this may be different in the case of, for example, "scheming plan" and the scenario attached to it on which *Othello's* plot builds. I suggest that one of the functions of our tendency to tell stories is that it allows language users to get accustomed to a wide set of such tools available in their cultural-linguistic community.

The question finally remaining now, is how all of this affects our evolutionary story. It is important to emphasise the qualification I have put forward in Chapter 1 once again: the material studied throughout this thesis and the arguments developed on the basis of it should not be taken as an attempt to "prove" a particular chronology of evolutionary events. However, given that the ways in which different scenarios of human cognitive evolution can be compared are necessarily limited (archaeology and comparative biology can tell us only so much in this domain), I have suggested to look at how

complex thoughtscapes are handled in the actual practice of human interaction. What exactly do we find in a play or novel renowned for dealing with wide arrays of perspectives? How do newspapers report on cases in which multiple intentional states are relevant? What can we learn from usage-based linguistic approaches as developed by Clark (1996) and Verhagen (2005)? How do participants seem to deal with experimental tasks forcing them to consider multiple intentional states at the same time? By investigating such questions one by one, a view has emerged that certainly can be said to be *more compatible* with one evolutionary scenario than with the other.

In this light I have suggested a "move" that is conceptually comparable to the one made by Shultz, Opie, and Atkinson (2011):[105] the route to a cultural-linguistic community should not be seen as one in which individuals developed increasingly profound mindreading abilities, going step-by-step from "A intends that B believes…" to "A intends that B believes that A intends…" to "A intends that B believes that A intends that B understands…", and so on, until addressees and speakers had arrived at fourth- and fifth-order intentionality respectively, which they needed to establish the sort of cognitive "interdependence" Scott-Phillips (2015) and Sperber (2000) consider to be a prerequisite for human ("ostensive-inferential") communication. Instead, the alternative supported in this thesis is that a form of shared intentionality *preceded* the capacity to deal with complex thoughtscapes, including the ability to work out in detail what A believes that B believes that A believes (…etc.). Roughly, instead of interaction in dyads growing more sophisticated, gradually including more "minds" and thereby eventually forming the basis for cultural-linguistic communities, I suggest that the presence of such a community was needed for increasingly sophisticated interaction in dyads.

At the end of Section 5.2.3 in Chapter 5 I have discussed the cognitive structure individuals need for participating in a system of interaction based on shared intentionality from a *synchronic* perspective, thus abstracting from the *diachronic* story of how such a system could emerge over time (see also the first

[105] Clearly, there are also important differences: their model concerns going from a situation in which individuals form no group at all to one in which there *is* a social group, whereas my suggested move concerns a situation in which there already is a group to one in which individuals within this group form a cultural-linguistic community.

note in Chapter 5). I have listed the following abilities: (i) distinguishing between individuals and recognising whether or not they belong to our social network and/or cultural-linguistic community (in order to identify the appropriate sort of common ground); (ii) keeping track of former interactions and accessing "records" of this in real time; (iii) applying the principle of self-other equivalence; and (iv) assuming a particular set of intentional states to be shared with one's interlocutor and singling out individual perspectives only if the context requires to do so. In order to put together the diachronic story it would be necessary to determine when and to which degree these abilities were available to our ancestors. I think that more research efforts should be invested in this in the near future;[106] however, having said that, I will end by giving it a provisional shot.

The abilities (i) and (ii) are in principle primate skills, found in our great-ape relatives to a degree of sophistication that I could imagine to be sufficient for getting started. In my view there is thus no reason that these would have been factors limiting the emergence of the sort of interaction system suggested here at some point in our evolutionary path after the divide from the other great-ape lineages (that is not to say that we have not become better at these at later stages). Ability (iii), applying the principle of self-other equivalence, is what I have discussed as the skill needed for taking part in the "Schelling mirror world" (Levinson, 2006) in Chapter 5. It is fair to assume that this ability requires a degree of mindreading beyond what our great-ape relatives can do, but there is no reason why it should involve higher-order intentionality: interlocutors have to reflect on their own and the other's perspective, and they would need to be capable of imagining that roles be reversed in order to consider contributions to the interaction from both sides. This involves working with two related, but not necessarily embedded, intentional states in a flexible way.

Fitting ability (iv) into the diachronic story first involves specifying the conditions under which it makes sense at all for individuals to enter an

[106] As mentioned in Chapter 5, Stolk's (2014) experimental work provides an important basis, but it comes with the unavoidable issue that the participants have already been through the process of growing up in a present-day cultural-linguistic community.

interaction event under the assumption that a set of intentional states is mutually shared. I suggest that an attempt to deal with this issue should at least include the elements of *cooperation* and *repeated interactions* between a somewhat stable set of individuals. In this way, it is of interest to individual participants to find an efficient solution for achieving a common goal. Roughly put, if repeated interactions lack common goals, there is no incentive for individuals to invest effort in coming to any solution; if common goals have to be achieved only occasionally there is little incentive for coming to an efficient solution. It is when trying to achieve a joint goal in an efficient way that it makes sense to assume a set of shared intentional states.

This then forms a context in which the emergence of conventional signals for singling out individual perspectives and negotiating deviations from a common ground makes sense—after all, entering an interaction event under the assumption that a particular set of intentional states is shared may form a good *starting point*, but if the exact same set is also necessarily the *end point* of the interaction, it has little practical use. With the availability of signals for negotiating how this set of shared intentional states should be updated, such usefulness increases. In this way, I suggest that a co-evolutionary loop was triggered in which increasingly profound negotiation signals led to wider applicability of the assumption of shared intentionality and vice versa, eventually producing the cultural-linguistic communities in which the sophisticated linguistic items I have argued to be crucial for both communication and processing of complex thoughtscapes could coagulate and get passed on from generation to generation. As soon as such items are available to members of a group, they introduce a sort of supra-individual order capable of orchestrating joint actions, saving individuals lots of negotiation costs including time, risk, and cognitive power.

In Chapter 1, Section 1.3.2, I have discussed the "bonding gap", the idea that our hominid ancestors needed to find new ways of bonding increasingly large social groups. Whereas our primate relatives, and presumably also the last common ancestors we share with them, use(d) time-consuming one-on-one activities (social grooming in particular), our ancestors must at some point have relied on more efficient alternatives such as laughing, dancing, and singing together. It has been suggested that language also played a key role, but

contributed to social bonding in a more indirect way through, for instance, gossiping or sharing jokes and stories. These activities (except maybe telling basic jokes) have been argued to require the availability of a fairly sophisticated form of language, and given the assumption that such a language demanded around five orders of intentionality to be in place, its arrival was positioned late in our evolutionary history. Although I agree that storytelling involves quite sophisticated language forms, I have argued that higher-order intentionality is not a prerequisite for this, and that the ability to deal with complex thoughtscapes is not so much inherent in individual cognitive capacities, but rather produced by an interplay between individual cognitive factors and phenomena that are part of language and narrative itself. This points towards a scenario in which linguistic and narrative abilities, cultural-linguistic communities, and increasingly sophisticated mindreading skills have co-evolved rather than one in which the latter are a prerequisite for the former to emerge.

To conclude, the key elements in my provisional version of the diachronic story of interaction based on shared intentionality that are *not* likely to have been present in the last common ancestor with our great-ape relatives, are thus the mindreading skills for operating in the Schelling mirror world and forms of cooperation among somewhat stable groups of individuals geared towards achieving common goals. Gamble, Gowlett, and Dunbar (2014: 140-146) convincingly argue on the basis of brain size, complexity of artefacts, and patterns of their socially mediated activities on landscapes that these elements were available to the makers of the famous Acheulean hand axes: *homo erectus* who lived in the Lower Palaeolithic from around 1.7 million years ago. Given my arguments in this final section there is clearly no reason to assume that our ancestors at that stage possessed fully-blown linguistic abilities, instantaneously involving such sophisticated features as gossiping and storytelling. However, I suggest that the capacity to deal with higher-order intentionality can be no reason for why the co-evolutionary loop described above could not have started in this era. From there, I suggest that our social brain includes an *individual* and a *collective, socio-cultural* dimension. Our evolution in increasingly complex social environments has not just put pressure on our individual brains to get bigger and more powerful (i.e. the classic social

brain hypothesis), but also on finding culturally transferable solutions for using these brains in an optimally efficient way. These solutions can today be found in the toolkit that language and narrative offer us for dealing with complex social situations and their underlying thoughtscapes.

I very much look forward to future arguments and evidence supporting or challenging my version of our story, thereby shedding more light on the process of which the outcome has been studied in this thesis: on the individual level we are lazy mindreaders who could do little intentional reasoning with our "bare" brains, but empowered with the toolkits available in our cultural-linguistic communities we appreciate Shakespeare and Woolf, project complex thoughtscapes when reading headlines, enjoy *Friends*-style practical jokes, or can even consider taking up a career as a double agent.