What makes us social and what does it tell us about mental disorders?

Uta Frith & Chris Frith

To cite this article: Uta Frith & Chris Frith (28 Jan 2024): What makes us social and what does it tell us about mental disorders?, Cognitive Neuropsychiatry, DOI: 10.1080/13546805.2024.2307958

To link to this article: https://doi.org/10.1080/13546805.2024.2307958
What makes us social and what does it tell us about mental disorders?

Uta Fritha and Chris Frithb,c

aInstitute of Cognitive Neuroscience, University College London, London, UK; bInstitute of Philosophy, School of Advanced Study, University of London, London, UK; cWellcome Centre for Human Neuroimaging, University College London, London, UK

ARTICLE HISTORY Received 28 December 2023; Accepted 16 January 2024

Our new book, “What Makes Us Social?” (Frith & Frith, 2023 published open access), started with four lectures on our joint topic of social cognitive neuroscience. We gave these lectures at the École Normale Supérieure in Paris in 2014 with the understanding that we would publish them as part of the Jean Nicod lecture series. We first planned for seven chapters. However, over the years this book grew to 20 chapters reflecting the rapid growth of the field. We were pleased to have been asked to comment on our book’s implications for mental disorders. We welcomed this task since our interest in social cognition stems from our involvement in research in autism and schizophrenia, conditions where problems with social communication and interaction loom large.

In the past, we have explored the role played by impairments in the ability to mentalise, that is, to spontaneously attribute mental states to self and others, also known as Theory of Mind (Apperly, 2012). Rather than observing the visible preferences of others to predict what another person is going to do next, mentalising allows inferences about the invisible intentions of others. This is a huge advantage in social interactions, and we argued that a deficiency would result in not comprehending why others act as they do. Thus a weakness or absence of spontaneous mentalising goes some way to explain the typical problems of autistic individuals with everyday reciprocal communication (Frith, 1989). An over-attri-

However, there are many more processes involved in social communication and interaction, such as alignment, imitation and the spontaneous formation of ingroups. So, it is right to ask whether any of these processes could allow us to penetrate deeper into psychiatric symptoms in social communication and interaction. There is plenty of room to develop hypotheses for understanding breakdown in social communication and test them with novel and exciting paradigms, for instance, using contagious laughter (Scott et al., 2014), over-imitation (Lyons et al., 2007) and audience effects (Hamilton & Lind, 2016).

CONTACT Uta Frith u.frith@ucl.ac.uk Institute of Cognitive Neuroscience, University College London, 17 Queen Square, London WC1N 3AR, UK

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.
A hierarchy of computation

To understand better the variety of processes, we used a framework that made sense in terms of evolution and explained what is unique in human social interactions. It covers three different realms: the physical world of objects, the biological world of agents and the mental world of ideas. Social cognition is important even in the world of objects. Like nearly all animals we learn about this world by observing and copying the behaviour of others, what they approach and what they avoid, constrained by the laws of physics and the organisation of their bodies. This is vicarious social learning, which saves animals from having to make their own, potentially fatal, mistakes. We tend to think that this level of processing is universal in social animals including humans, and unlikely to help explain psychiatric disorders.

As we move from objects to agents and ideas, the cognitive processes needed for successful interactions become more complex. Agents can learn even more successfully what to approach and what to avoid, when they automatically assume that agents act on the basis of their goals. However, different agents can have different preferences, and there may be no need to approach what they approach or to avoid what they avoid. These processes too are unlikely to yield explanations of mental disorders.

In the world of ideas, which is probably only inhabited by humans, we automatically assume that people act on the basis of their beliefs and intentions. Inferences about their mental states allow a far better way to predict what they are going to do than merely observing what goals they approach or avoid. Here we can share beliefs about the world and work together to make them more accurate. We can also use deception to manipulate the beliefs of others to obtain advantages. This is why in the book we identify mentalising as the dark heart of social cognition. But the world of ideas is much richer. It contains the interface between our mind and other minds. This is where even more complex social processes are situated, and these complex social processes have clear implications for psychiatric disorders.

For simplicity, in Figure 1 we only represent two levels, when there are likely many more, each requiring a mechanism that handles social information as well as a metacognitive monitoring mechanism. Processes below the bold line are deeply unconscious and shared with all other social animals. The conscious processes at the top of the information processing hierarchy may only occur in humans. We are certainly the only creatures who can talk to each other about our inner, subjective experiences. We can, for example, check their reality by asking another person “Do you hear a strange buzzing noise or is it just in my head?”.

We refer to this sort of checking as metacognition. This is not reserved only for conscious thought but occurs at all the levels of the hierarchy. What does metacognition do for us? It serves as a monitor, and it is also a means for communication between the different levels. It produces alarm signals when something has gone wrong and sends them up to higher levels. From there signals can be sent down to alter the functioning of the lower level (Norman & Shallice, 1986), for example, by issuing the command to slow down.

The prediction error, which is critical for learning, is an important metacognitive signal. It indicates that the cognitive mechanism that is being monitored is not working as expected. Prediction errors play a fundamental role in learning (Schultz
et al., 1997). There is evidence that disruptions in the handling of prediction errors may be associated with psychotic delusions (Corlett et al., 2007).

Another important metacognitive parameter is “precision” (Yon & Frith, 2021). Precision is the same as the reliability or variance of a signal and is crucial when distinguishing signals from noise. If the precision of a signal is high, then we should take notice. Conversely, signals that have low precision should not be given the same weight. For example, we ignore a burglar alarm that can easily be set off by the fluttering of a moth. Precision at the bottom of our computational hierarchy concerns the reliability of sensory channels, such as vision and touch (e.g., vision is more reliable than touch, Ernst & Banks, 2002). This is an example of implicit metacognition. We are not aware that this reliability is being considered when we combine vision and touch. Precision at the top of the hierarchy concerns the strength of consciously held beliefs, an example of explicit metacognition. As an example we can take a current credo of cognitive neuroscience: “Free energy is the key to understanding how the brain works”. Here we are uncomfortably aware that the strength of this belief may be undermined by our lack of precision concerning the concept of free energy.

**Explicit metacognition and culture**

But what has metacognition got to do with social cognition? This becomes apparent when we look at the very top of the processing hierarchy. Here is where information gets from one brain into another! In fact, the interface at this top level is entirely social, and the
processes involved are processes of social cognition. Once we enter the world of ideas we start to interact with the ideas of others.

Explicit metacognition effectively adds another layer to our information processing system. This higher-level layer concerns other people and our culture more generally. It enables culture to influence individuals and enables individuals to influence culture (Sperber, 1996). We can ask others for advice when we realise that things are going wrong, and the advice we are given will change our behaviour. Even though social interactions take place at the top of the hierarchy their effects can reach all the way down. We marvel at the fact that simple instructions from others can create habits which operate at the lowest level of the hierarchy. For example, in the UK we teach young children to look to the right before crossing the road. This creates a habit. Like other habits, it is difficult to overcome. Thus, when we go abroad, we must consciously suppress this habit to look to the left.

There is wide agreement that culture shapes us as individuals at least as much and possibly more than our innate predispositions and personal history. We consider teaching key to this process. There are reasons to believe that direct active teaching (Kline, 2015) is a uniquely human behaviour that fast-tracks transmission of information. Culture affects the way we perceive the world, so that any sensory differences between us are minimised. For instance, we pick up rare names for shades of colour (e.g., sepia, magenta, ochre) and fine tune our vocabulary and our perceptions accordingly. People who are colour blind or have synaesthesia often do not realise until adulthood that they see the world differently from other people (Bradley, 1970).

These examples suggest that our interactions with each other and culture more generally, smooth out the differences between us. We see the world in the same way and acquire common knowledge about how the world works. We acquire common folk psychological theories about the causes of behaviour. For example, in the second half of the twentieth century psychoanalytic ideas, e.g., the Oedipus conflict, became ubiquitous in fiction (Haslam & Ye, 2019). Currently, the idea that our behaviour is determined by hidden states, e.g., unconscious perceptions and memories (Godfrey-Smith, 2005), seems only common sense. These ideas are not only applied to understand the behaviour of others, but they also become norms for determining our behaviour (McGeer, 2007). As a result, even if these theories are wrong, their incorporation into common knowledge makes social interactions easier to negotiate.

The top-down effects of culture on perceptions, actions and beliefs have an important role in enhancing social cohesion. For example, we learn about the concepts of equity and fairness. Most of us, when perceiving unfair behaviour, whether to ourselves or others, experience an immediate feeling of unpleasantness (Haruno & Frith, 2010). It follows that behaving fairly becomes a habitual action which we don’t have to think about (Rand et al., 2012). However, these feelings of fairness can be overcome by new, culturally formed ideas about equity. For example, “trickledown economics” and “effective altruism” are ideas designed to make us believe that becoming much richer than other people is good for society and therefore fair (MacAskill, 2015).

We learn about the world and about ourselves, through interactions. Most successful learning requires many, many interactions. For example, we become very skilled at reaching and grasping since this is something we do very frequently every day. Through such extensive learning we acquire the appropriate prior expectations for
successful interactions. But, particularly in the social domain, there are many kinds of interaction which do not occur very frequently (e.g., meeting the King). As a result, we do not learn how to behave in such contexts, and our prior expectations remain stuck in a non-optimal state. In these circumstances, learning from others becomes crucial (e.g., Debrett’s Etiquette).

We benefit from the past interactions of many other individuals, and the resulting wisdom (or sometimes folly) has become embedded in our culture. Instructions from others can rapidly change our priors, even when we have no direct experience of the task being learned. For example, in fear conditioning we can gradually learn by trial and error that a blue square on the screen will be followed by a nasty shock. Eventually the sight of the blue square will elicit a fear response. This effect depends upon Pavlovian conditioning. Initially the blue square has a neutral prior value, but this value is slowly updated as we experience the shocks. However, we can also learn from instruction (e.g., 

**Implications for mental disorders**

Will our approach to social cognition help us to better understand mental disorders?

We presume that new ideas might help to explain psychiatric symptoms rather than current diagnostic categories. We would place our bets on problems at the level of explicit metacognition, and we will give an example. We also hope that the cultural interface at the top level of the hierarchy can throw light on the continuous changes in diagnostic categories. We explore this further in another example.

*Explaining symptoms as problems at the top of the hierarchy.* People with psychosis are typically described as having lost contact with reality. But it would be more accurate to say that they have lost contact with our reality, i.e., the world that the rest of us experience and understand in our culture. In our framework, this would imply that something has gone wrong with explicit metacognition. The problem applies to both aspects of metacognition: monitoring (signals going upwards and/or outside to others) and control (signals going downwards and/or initiated by others). At the interface with other minds, patients have difficulty in reporting their subjective experiences in ways that others can understand and so cannot be properly monitored. Likewise, the reports or instructions by others no longer have control on their cognitive functioning. There is a loss of cultural constraint, so that the patient’s model of the world inevitably deviates further and further from the model shared by everyone else. From here it is only a small step to frame their ideas as “mad” or “bizarre”.

There has been much interest in using a Bayesian (predictive processing) approach to the understanding of psychotic symptoms (e.g., Petrovic & Sterzer, 2023). Using this approach, the resistance of delusions (high level beliefs) to updating by evidence can be explained by assuming that the priors associated with the delusion are too sticky and incoming evidence is ignored. Perhaps the priors are too precise (i.e., wrongly trusted as reliable). However, this explanation does not tell us why the level of precision, and hence the high weight given to bizarre beliefs, has become so abnormally high.

One suggestion is that, even in the normal case, high-level priors are sticky and don’t yield easily to demands for change. These priors are sticky when there is not enough
directly relevant experience without which they cannot be updated. This is not normally a problem, because interaction with others can change these same priors rapidly. The effectiveness of instructions is proof that we can trade direct experience with indirect experience as conveyed by social interaction.

In contrast, patients with bizarre delusions cannot be talked out of them. We propose that this is a problem with explicit metacognition, and more precisely, with the mechanism that normally enables social interactions and instructions to change beliefs. Is this because the precision associated with the belief is too high, or is it because the advice from others is constantly downweighed?

The phenomenon of *folie a deux* provides a striking example of this difference between standard metacognition and psychotic metacognition. The psychotic member of the pair develops delusions that are increasingly at odds with the surrounding culture, whereas non-psychotic partners align their beliefs and attitudes with their psychotic partner. In this interaction the influence on high order beliefs travels only one way. Once the non-psychotic member of the pair has been separated from their partner, their delusions can often be reduced through rational argument (Langdon, 2013).

**Explaining changes in diagnostic categories as effects of cultural influence.** In our book, we suggest that cumulative culture emerges from social interactions via explicit metacognition. Through these interactions we modify our concepts and continuously update our shared models of the world. Medical and psychiatric concepts are not immune to this process. Over the last few decades, we have seen this happening in the case of the concept of autism. In the 1960s, when we were studying for our PhDs, autism was a rare disorder (a prevalence of ∼0.5 per 1000) associated with intellectual disability. Today the median prevalence worldwide is about 10 per 1000 (Zeidan et al., 2022). A dramatic increase over the last 20 years (Russell et al., 2022) reflects a striking change in our concept of autism.

One cause of this change was the idea, associated with Hans Asperger (Tantam, 1988), that autism could be found in individuals of normal or high intelligence. People were fascinated by this idea. What would such a person be like? Asperger had described “little professors”, children who could talk fluently about their special interest without noting the signs of boredom in their listeners. We all seemed to know adults like that, often real and sometimes eminent professors. The concept of autism was beginning to escape from the clinic and enter the wider world.

At around the same time (1988), the film “Rainman” was released and raised awareness of autism in adults. The public was introduced to an autistic protagonist who was socially inept, rigid and unable to live by himself, but at the same time a savant, capable of amazing feats of memory and calculation. Savant skills became part of the concept of autism even though such skills are only found in ∼10% of cases (Treffert, 2009). Since then, similarly rigid, but highly intelligent characters are frequently represented in the media. Currently they are an icon as detectives who defy convention and use their special skills to reach the truth (e.g., Saga Norén in “The Bridge” and Astrid Nielsen in “Murder in Paris”). These characters are typically much less socially impaired than the character in Rainman. But their autistic traits are instantly recognisable to everyone.

Today autism is no longer a “spectrum of disorders”, as originally proposed by Lorna Wing (1997), but a “neurodiverse” spectrum that shades into “neurotypical”. In terms of this new spectrum, we can all be a little bit autistic. And there are many sites on the internet which provide tests to reveal how autistic you might be. This concept of autism is very
different from the one we understood 60 years ago (Happé & Frith, 2020). To us this new concept seems more like a personality dimension (Camp, 2023). But this is a rather special kind of personality. Through the effects of human cumulative culture, labelled “looping” by Ian Hacking, autism has become a new “way to be a person, to experience oneself, to live in society” (Hacking, 2006).

Given this view, the term disorder is no longer appropriate, and this is problematic for clinical research and practice. Indeed, this problem has begun to be discussed (Fombonne, 2023) (Mottron & Gagnon, 2023), but it is not yet clear how it will be resolved. Perhaps new diagnostic categories will emerge.

*Mental disorders are fundamentally social.* The problem with the diagnosis of autism is not novel. The same problem has long affected diagnoses associated with those most common symptoms of mental disorders, anxiety and depression. On the one hand, there are severe cases of anxiety and depression that can be extremely debilitating and require treatment. On the other hand, there are dimensions of anxiety and depression. Our positions on these dimensions are continually varying, and all of us, at some point, will experience uncomfortable extremes. But we will recover, from these extremes, typically through the kindness of friends.

The need for kindness and the need for friends relates to the major theme of our book. We are all far more social than we realise. All animals learn from others and together with others, but we humans are even more socially embedded. With our large communities and our emphasis on division of labour, we are more like ants and bees than like other primates. We are also embedded in a cultural world of ideas. And yet we persist in thinking of ourselves as self-sufficient individuals, free to do as we please.

During the pandemic and lockdown, there was a dramatic increase in self-assessed depression and anxiety (Dettmann et al., 2022). It is likely that a major cause of this increase was the loss of direct contact between people. This loss was made more damaging by the threat created by the virus. Our natural response to threat is to seek more contact with others (Dezecache et al., 2020), but we had to inhibit this urge. It seems that an escalation of anxiety and depression were inevitable. This line of reasoning is only confirmed by the heightened enjoyment of each other’s company and subsequent mood elevation, we individually experienced after the lockdown.

The major message of our book is that almost all cognition is social cognition. Likewise, perhaps almost all mental disorders are disorders of social cognition.

**Disclosure statement**

No potential conflict of interest was reported by the author(s).

**Funding**

The author(s) reported there is no funding associated with the work featured in this article.

**References**


Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature, 415*(6870), 429–433. [https://doi.org/10.1038/415429a](https://doi.org/10.1038/415429a)


Kline, M. A. (2015). How to learn about teaching: An evolutionary framework for the study of teaching behavior in humans and other animals. *Behavioral and Brain Sciences, 38*, e31. [https://doi.org/10.1017/s0140525x14000090](https://doi.org/10.1017/s0140525x14000090)


