Inequalities in mental health: predictive processing and social life.

Michael P. Kelly1, Carol Brayne1\*, Ann Louise Kinmonth1, Natasha Kriznik2, John Ford1, and Paul C. Fletcher3.

1. Department of Public Health and Primary Care, University of Cambridge, East Forvie Building, Forvie Site, Cambridge. CB2 0SR. UK.

2. Birmingham City University, Birmingham B5 5JU. UK.

3. Department of Psychiatry, University of Cambridge, Douglas House, 18b Trumpington Road, Cambridge. CB2 8AH. UK.

\*Corresponding author Department of Public Health and Primary Care, University of Cambridge, East Forvie Building, Forvie Site, Cambridge. CB2 0SR. UK. cb105@medschl.cam.ac.uk +44 01223 330300

**Abstract.**

Purpose.

The paper applies recent conceptualisations of predictive processing to the understanding of inequalities in mental health.

Recent findings.

Social neuroscience has developed important ideas about the way the brain models the external world and how the interface between cognitive and cultural processes interact. These recall earlier ideas from cybernetics and sociology. These ideas may applied to understanding some of the dynamics leading to the pattering of mental health problems in populations.

Summary.

The implications for practice are the way these ideas help illuminate thinking and acting, and how they are anchored in the social world.

**Keywords.**

Social neuroscience. Predictive processing. Health inequalities.

**Acknowledgements.**

Beyond the authors, there was no additional assistance provided in preparing this paper.

No funding was received in support of preparing this manuscript.

The authors have no conflicts of interest to declare.

**Introduction.**

There is strong evidence of relationships between economic, environmental and health disparities and all forms of well-being and mental illness [1]. The conditions in which people are born, grow, live, work and die have profound effects on their mental health. We examine the social mechanisms by which the brain makes sense of these conditions [2, 3].

One cornerstone of good mental health is a sense of being at one with the world [4, 5]. If our understanding of the external world actually aligns with that external world, it is possible to act in, and on that world, with a reasonable degree of confidence, and confidence that the external world will respond in ways congruent with our expectations. Antonovsky called this having a ‘sense of coherence’ [4, 5]. The problem is that for many, there is a disjunction, more or less acute, between their own understanding of the external world and the way the world responds to them. This can take many forms, some of which are experienced and defined as mental ill health. However, the degree of disjunction is not randomly distributed in the population. It is patterned because it is associated with social position, access to resources and the ability to control one’s own lifeworld. These in turn are strongly associated with socio–economic position.

We explore the processes involved by applying a framework that is increasingly influential in neuroscience - predictive processing. Predictive processing is a theory of brain function in which the brain is constantly generating and updating a model of the world, a model that shapes predictions about the sensory inputs from that world. These predictions are theories with inherent biases and assumptions. Mismatches between the predicted and the actual inputs, prediction errors, constitute an updating signal that informs future predictions. Thus, alluding to Antonovsky’s ‘sense of coherence’ above, predictive processing frames the activity of the brain as engaged in a quest for coherence, by optimising predictions to minimise mismatches. The theory highlights important links between neuroscience and social and cultural processes [6]. This framework provides potential insights to understanding inequalities in mental health.

**Predictive processing: an overview**.

Predictive processing has roots in a number of scientific fields, common to which is the acknowledgement that, as ‘agents’, (i.e. autonomous entities that interact with the world), humans are not simply passive recipients of sense data. Rather, humans actively participate in constructing a model of the world and updating this model through an iterative process of making predictions, sampling data and rejecting or taking actions. By sampling the world, people update their model to align more closely with it. By acting on the world, they alter it, to align more closely with their model.

The notion that human brains embody a model of the world was established in the field of cybernetics [7]. It was articulated in a set of laws stating that, to be successful, the agent must recapitulate in its structure and function the features of the world that it inhabits and that its capacity to regulate that world demands a repertoire of responses that mirror the influences that the world may have upon it. Again, we see a strong echo with Antonovsky’s notion of a ‘sense of coherence’ emerging from the interactive relationship between the inner model and the external world.

Recognising these attributes, we confront the idea that comprehending ourselves and our brain function demands an understanding of our environment and the varying impacts that this environment can have on us and vice versa. The biopsychosocial perspective assigns due weight to the individual aspects of and explanation. A synthesis of the interactions between biological, psychological and social phenomena emphasises dynamic and evolving relationship between the agent and its world. In turn this demands a closer inspection of how they interact, and then come together to shape us. Such an emphasis supports a developmental view, scrutinising the social and cultural worlds into which the infant is born and in which the child develops and operates through adolescence into adulthood. This view promises a more comprehensive consideration of socially determined mental health and links directly to sociological theories concerned with action, structure and social practices relating to inequality.

**Predictive processing is not a uniform process but a dynamic reflective one.**

Amongst many things, the external social environment in which people live and work is characterised by differential access to power and resources. It is the setting for discrimination and disadvantage, from the subtle to the explicit, as well as more benign experiences. For good and ill the external environment, and experience of it, are the basis of the way that individuals think about and perceive the world, and the way that they seek to act on and to operate in it. The individual’s constructed models are by definition subjective. They are therefore highly heterogeneous. They will vary in their alignment with the objective world *and* the world as others see it. The interaction between an agent and their world is partly determined by the nature of the agent and partly by the features of the world. Two agents might inhabit the same environment but have very different *Umwelten* depending on their structural properties but also upon the model that they brought to that world or that they project onto that world [8].

Some models of the world held by an individual brain may be optimal in terms of surviving in the here and now, but, in the longer run, may prove dysfunctional in attempts to operate in the broader environment and in the eyes of that wider external world. As a simple example, it has been found that substance dependence is associated with “poor” performance in temporal discounting tasks, which require a capacity to resist immediate reward in exchange for a greater, delayed reward [9]. But of course if someone inhabits, or has expectations of, a world of uncertainty and volatility, then temporal discounting is an optimal strategy, making the individual more successful in that immediate environment. Everyone has to bridge the potential gap between the world as they think and hope it is, and the way the external world is found to be. For some the gap is not that great and is easily managed, for others the gap is much more difficult to close.

Trauma, disease, drugs, or alcohol may permanently or temporarily perturb the ability to do predictive processing (i.e. to integrate expected and actual inputs). This may push the process in particular directions (e.g. create the expectation of danger based on long experience even when the current environment is safe and does not require vigilance or defensive strategies). The agent may sow patterns of behaviour, thought, action and affect, which impair ability to model the world and to achieve equilibrium in social interaction.

**Intersubjectivity.**

Alfred Schutz writing in the mid-twentieth century argued that intersubjectivity was a core task, which we all have to engage with continuously [10,11]. What he meant is that we have subjective feelings and thoughts. Some of these we explicitly articulate to ourselves, others are much more implicit and what he called “taken for granted”, what today psychologists sometimes refer to as automatic processes [12]. The problem is that these ideas, *qua* ideas, are not directly observable to others and if we wish to share our thoughts and feelings we have to talk to others, or signal our subjectivity by our actions, deeds, deportment and style. And of course we may inadvertently or deliberately try to mislead others in our words and actions. Others interpret what they see and hear us do and say, and read off from what we say and do, what they think we are thinking and feeling, or what they expect us to be thinking and feeling – their interpretation of our subjectivity. They then act on that basis.

This happens in the here and now, but is also the product of experience across the life-course. Certain habitual ways of processing what we see and hear others doing, and they us, arise and are learned in the social practices in which people engage [13, 14]. And these are of course highly variable. Notwithstanding this heterogeneity, we still have to engage with others, and fathom out their subjectivity in order to interact with them. A shared subjectivity –intersubjectivity- Schutz suggested, is the very thing, which makes human interaction possible. This is what predictive processing facilitates. Without a shared understanding with others, our ability to empathise or to have a conversation, for example would be very restricted.

The person who has suffered early neglect/trauma/abuse may tend towards a world model that anticipates bad intentions in others [15]. The predictive processing model emphasises that they will act on and sample the world in such a way as to fulfil this prediction (to minimise prediction error). The person inadvertently creates situations that fit with this model. The transfer of the model from one context (a dangerous one in which they must be hyper-vigilant and untrusting) to another (where threat is diminished and the optimal strategy is trust) can lead to an emergence of unfounded paranoia and profound socially awkward interaction in more benign conditions.

Schutz also noted that our understandings of others’ subjectivity, and theirs of ours, is always imperfect, sometimes massively so. We have to update our models. Most humans as it turns out, manage this ambiguity and uncertainty reasonably well most of the time, and indeed are highly skilled in making allowances for, and getting past *faux pas*, mistakes, and misunderstandings [16,17]. We act on the basis that we assume that what is going on in other peoples’ heads most of the time, is broadly comparable to what is going on in our own minds, even though we also know it is not in any sense exactly the same. In the contemporary parlance of social neuroscience, we adjust our predictions and remodel our view of the external world.

**Ideal types and typifications.**

We are, in the language of social neuroscience, continually engaged in a process of testing our theories - our predictions about other people and their actions and words, and refining our predictions based on the emerging evidence. Predictive processing emphasises that the models people use are built by testing new information as it becomes available. The models also use prior knowledge and understanding based on previous experience. Sociology is interested in priors in this sense, and some theorists have used the terms ideal types and typifications to describe them [3, 10, 11].

Ideal types define macro ideas about how the world works. Typifications are more detailed conceptions [6]. Ideal types are broad classificatory frameworks that are ideal in the sense that they do not exist – they are ideas - but they provide ways to think about very generally, all kinds of phenomena. So concepts such as religion, democracy, war, decency, love, bureaucracy, power, gender, ethnicity, trust and many more, are rough high-level guides as ideal types to what is out there in the world. They lack detail and have to be general enough to deal both with things we might be familiar with from our day to day lives, and to things we have never encountered directly or indeed never existed, but have learned or heard about vicariously. Such things include the Court of King Arthur, the Napoleonic War, American slavery before the Civil War, women not having the vote before the 1920s in the UK, for instance. Typifications are much more detailed and often grounded in experience. Once again, they are not precisely accurate images of the world, but they provide frameworks, which are good enough for most people most of the time. Typification are a Bayesian predictive tool [18]. In social neuroscience, we can think of these typifications as the building blocks of predictive processing, which have to be plastic enough to adjust to new inputs. The typifications and ideal types are the stuff that makes up the brain’s model of the world. As they are refined and sharpened, they allow the brain’s model to be updated. Both ideal types and typifications originate in the social world [3].

One of the features of the predictive processing framework (Active Inference) suggests that people do not just predict, sample and update their priors but act in order to fulfil predictions, or to minimise prediction errors. They select and generate an action based on a prediction error between a predicted state and a current state. If these states differ, they generate a prediction error that they are motivated to minimise. This is the way that typifications may work because as well as providing the classification networks that govern predictions/priors in the social domain, they will help shape the way the individual strives to act in ways that reinforce or confirm, or struggle to reject, the typification.

Problems arise when the remodelling process fails to create updated models that are congruent with the external world and then hence the next set of inputs that brain faces. This happens when the typifications are not flexible enough to capture the nature of the complex external reality and may be a consequence of capacity, learning, or prior experiences, such as the systematic experience of social exclusion and discrimination. If models are not updated and typifications are unsuited to the environment in which the human is currently operating, then that person finds themselves maladapted. All humans experience maladaptations and disjunction, but mostly they are able to navigate their way away from environments where they don’t fit, or they change their predictions about the world during processes of socialization into the new social practices they encounter.

However, maladaptation may be chronic and unresolved, or unresolvable. The issue becomes chronic rather than transient for many possible reasons. The person may not possess the ideal types, the ability to use typifications appropriately or helpfully, the resources to escape, or where in their efforts to confirm their typifications they collide with an obdurate alternative reality. Some ideal types and typifications will be culturally determined. So if the habitual way of seeing the world say in Native American society, is at odds with modern western cultural norms of the USA, the ability to switch types and typification may not be impossible, but can be tricky. Even in societies where there is apparent cultural homogeneity, in reality there is considerable diversity and therefore models will be heterogeneous. Societies like the US and UK are enormously variegated along class, gender, ethnic, geographical, historical, and age lines. In addition, these axes of social differentiation and their intersections make the possibilities of cultural ambiguity and misunderstanding potentially very great [19]. And where there is a dominant narrative linked to power and social exclusion, the capacity for the diversity, and its systems of ideal types and typifications to be experienced as legitimate and valued, is all the more problematic.

**Self and Identity**

There is a potential for a disjunction between self and identity. The self is the thinking subject, who knows that they are thinking, who can reflect and act on the external world and is aware of their capacities for reflexivity and agency. The self is the conscious awareness of the process of actively interpreting external and internal information and stimuli. Self operates by drawing upon its stocks of knowledge and its ideal types and typifications in so doing it is able to remodel the world through predictive processing.

The self has two elements, situational and substantial [20]. The situational self is the set of ideas that the self is engaging with in the immediate here and now. They relate to specific tasks or encounters of the moment, playing transitory roles of self, like customer, passer-by, or television viewer. The self has also a substantive dimension, as for example being a woman, a parent, a catholic, a mother – roles, which exist across time and place. For the situational self, typifications have to be agile and transferable, for the substantial self, longer lasting and concrete. For the situational self, there will be recourse to context - dependent predictions based on broad ideal types, which frame certain expectations (I’m in a forest so I’m more likely to see a bird than a fish). For the substantial self, the types and typifications are better defined as expectations that are embedded in the social roles and act automatically and ineluctably on all information. Appealing to a predictive processing framework, we might draw parallels between this distinction of the situational self, versus the substantial self and a distinction recently made between two different forms of prediction: those expectations that are context-dependent and those constraints that seem to be embedded in the structure of the agent [21].

But although the self has its own sense of who and what they are as a person, it does not follow that others will respond to that presentation of self in a way that is congruent with the presentation. The way others view self is called identity [22]. Where selves and identities are congruent and the external world acknowledges the person to be what they themselves think themselves to be, we have equilibrium. Where we have incongruence, we have disequilibrium, as people seek to be presenting themselves in a particular way, but others or society define them differently, in possibly stigmatising or discriminatory ways. If others deny the typifications that someone uses in interaction, the individual will find themselves in degrees of conflict with the external world on a more or less continuous basis. Alternatively, they will have to internalise the external definition of their identity into their conceptions of self. Predictive processing here faces either chronic and recurring problems of ambiguity or the need to undergo a paradigm shift in terms of self-conception.

Everybody experiences these disjunctions and everybody has to manage the minor disjunctions of everyday life. However, when chronic, disjunction is the seed of low self-image and self-esteem, loss of efficacy, and ultimately of potentially serious mental health problems. The same dynamic between self and identity will come into play where the internal self is significantly out of line with the structure of the social world because of traumatic brain damage or other causes of malfunction. The ideas then articulated by self, the order in which they are vocalised, the signals self receives from the external world mean that misalignment with the external world is inevitable [23]. The social world can provide an identity for this misalignment; gifted, second sighted, spiritual or mentally ill or demented [24]. The latter significantly reduces the power of the person to engage with the external world in a way that will change that new equilibrium.

**Effects of unequal societies**

These phenomena speak directly to the nature of predictive processing and inequality. They originate in social processes and structures and for many, those social processes provide a perfect storm where social disadvantage, overlays discrimination, overlays stigmatizing identity construction, to the point where it becomes very difficult either to change the brain’s model of the external world, or for the individual to acquire the predictive processing skills which will help them to change. Under such circumstances, harking back to Antonovsky, the only recourse to a sense of coherence is to act on the world in such a way that one’s expectations are more likely to be fulfilled. Perhaps it becomes more acceptable to reconcile oneself to unpleasant predictions that are fulfilled than to live in a setting of continual mismatch [5, 25].

Power, resources, wealth and education, are not distributed randomly in the population, they are highly patterned and the experiences of them is recursive and reinforcing. Consequently the predictive programming for all of us, for most a little, but for some people profoundly, are affected by the nature of society. Societies that are unequal generate dysfunctional predictive processing which fuels mental ill health for some members of that society.

**Conclusion.**

Both culture and cognition are supra-individual. It may be thought to be obvious that cognition takes place within an individual mind, but as that process takes place, the brain is engaged with the external world and is being affected by it continuously. The things with which it is engaged, are broadly cultural and symbolic, but some of those things are the handmaidens of class, power, privilege, poverty, racism, disadvantage and social exclusion. With mental illness, we need to note that most treatment involves individual interventions, but the very building blocks of the thought processes they seek to intervene on, and to help with, are themselves part of the social realm not individual.

3403

**Summary.**

* The brain operates with a model of the external world and builds that model by predictive processing.
* Predictive processing is an interactive social and cognitive process.
* Alignments and disjunctions between the brain’s model of the world and the worlds it is are the basis of good or poor mental health.
* The social word is profoundly unequal.
* Unresolved disjunctions between the brain’s model and the external world are disproportionately found in the more disadvantaged members of society.

**References.**

1. McDaid, Shari & Kousoulis Antonis (2020) Tackling social inequalities to reduce mental health problems: How everyone can flourish equally, London: Mental Health Foundation. <https://www.mentalhealth.org.uk/sites/default/files/MHF-tackling-inequalities-report_WEB.pdf>
2. Marmot, Michael (2010) Fair Society, Healthy Lives: Strategic Review of Health Inequalities in England post 2010, London: UCL. <http://www.ucl.ac.uk/gheg/marmotreview/Documents/finalreport>
3. \*Kelly, M.P., Kriznik, N.M., Kinmonth, A-L., Fletcher, P.C. (2019) The brain, self and society: a social-neuroscience model of predictive processing. Social Neuroscience, 14(3): 266-276. <https://www.tandfonline.com/doi/full/10.1080/17470919.2018.1471003>

An outline of the theory of predictive processing and its social origins.

1. Antonovsky, Aaron (1985) Health Stress and Coping, San Francisco: Jossey Bass.
2. Antonovsky, Aaron (1987) Unraveling the Mystery of Health: How People Manage Stress and Stay Well, San Francisco: Jossey Bass.
3. Clark, A. (2016) Surfing Uncertainty: Prediction, Action and the Embodied Mind, Oxford: Oxford University Press.
4. Conant, R.C. & Ashby, W.R. (1970) Every good regulator of a system must be a model of that system, Int. J. Systems Sci., 1970, 1: 89-97.
5. von Uexküll, Jakob (1957,1934) A stroll through the world of animals and men: a picture book of invisible world, in Claire H. Schiller, editor and translator, Instinctive Behavior: The Development of a Modern Concept, introduction by Karl S. Lashley, New York: International Universities Press. pp 5-80. First published 1934.
6. Robles, Elias. Huang, B. Emma. Simpson, Pippa M. Donald E. McMillan, Donald E. (2011) Delay discounting, impulsiveness, and addiction severity in opioid-dependent patients. Journal of Substance Abuse Treatment; 41 (4): 354-362. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3192942/>
7. Schutz, A. (1967) The Phenomenology of the Social World, trans G. Walsh & F. Lehnert, intro by G. Walsh, Evanston Ill: North Western University Press.
8. Schutz, A. (1970) On Phenomenology and Social relations: Selected Writings, ed and intro H.R. Wagner, Chicago: Chicago University Press.
9. Strack, F, Deutsch R. (2004) Reflective and impulsive determinants of social behavior. Personality and Social Psychology Review; 8:220-47.
10. Bourdieu, Pierre (1990) The Logic of Practice, Cambridge: Polity. Translated by Richard Nice, First published in French as le sens pratique, 1980, Les Editions de Minuit.
11. Bourdieu, Pierre (2000) Pascalian Meditations, translated by Richard Nice, Cambridge: Polity. First published in France as Meditations Pascalienne, Editions du Seuil, 1997.
12. Griffin JD, Fletcher PC. Predictive Processing, Source Monitoring, and Psychosis. Annu Rev Clin Psychol. 2017 May 8; 13:265-289.
13. Goffman, E. (1969) The Presentation of Self in Everyday Life, Harmondsworth, Middlesex: Penguin Books.
14. Garfinkel, H. (1967) Studies in Ethnomethodology, New Jersey: Prentice Hall. Cambridge: Polity 1984.
15. Seth, A. K. (2015). The Cybernetic Bayesian Brain - From Interoceptive Inference to Sensorimotor Contingencies, In T. Metzinger & J. M. Windt (Eds). Open MIND: 35(T). Frankfurt am Main: MIND Group. doi: 10.15502/9783958570108
16. Kelly, M.P. (2010) The axes of social differentiation and the evidence base on health equity. Journal of the Royal Society of Medicine, 103: 266-72, DOI .1258/jrsm.2010.100005
17. Kelly, M.P. (1992) Self, identity and radical surgery, Sociology of Health and Illness, 14: 390 – 415.
18. \*Teufel, Christoph and Fletcher, Paul C.(2020) Forms of prediction in the

nervous system, Nature Reviews Neuroscience. https://doi.org/10.1038/s41583-020-0275-5

<https://www.nature.com/articles/s41583-020-0275-5>

An elaboration and update of predictive processing.

1. Ball, D.W. (1972) Self and identity in the context of deviance: the case of criminal abortion, in Scott, R.A. & Douglas, J.D. (eds) Theoretical Perspectives on Deviance, New York: Basic Books.
2. Damasio, A. (1994) Descartes’ Error: Emotion, Reason and the Human Brain, New York: Putnam. Vintage Books: London 2006.
3. Meissner, W.W. (1978) The Paranoid Process, New York: Jason Aronson.
4. Geyer, Siegfried (1997) Some conceptual considerations on the sense of coherence, Social Science and Medicine; 44(12): 1771-1779. <https://reader.elsevier.com/reader/sd/pii/S0277953696002869?token=87B4E060628FE0B095A7F2E8B5872213A8FCB0724B24EB3260546AE29B0CF5834207CA1D76289B8E3FDB7A7A42575AC1>