



An active inference account of protective behaviours during the COVID-19 pandemic

Hugo Bottemanne^{1,2,3} · Karl J. Friston⁴

Accepted: 22 August 2021 / Published online: 15 October 2021
© The Psychonomic Society, Inc. 2021

Abstract

Newly emerging infectious diseases, such as the coronavirus (COVID-19), create new challenges for public healthcare systems. Before effective treatments, countering the spread of these infections depends on mitigating, protective behaviours such as social distancing, respecting lockdown, wearing masks, frequent handwashing, travel restrictions, and vaccine acceptance. Previous work has shown that the enacting protective behaviours depends on beliefs about individual vulnerability, threat severity, and one's ability to engage in such protective actions. However, little is known about the genesis of these beliefs in response to an infectious disease epidemic, and the cognitive mechanisms that may link these beliefs to decision making. Active inference (AI) is a recent approach to behavioural modelling that integrates embodied perception, action, belief updating, and decision making. This approach provides a framework to understand the behaviour of agents in situations that require planning under uncertainty. It assumes that the brain infers the hidden states that cause sensations, predicts the perceptual feedback produced by adaptive actions, and chooses actions that minimize expected surprise in the future. In this paper, we present a computational account describing how individuals update their beliefs about the risks and thereby commit to protective behaviours. We show how perceived risks, beliefs about future states, sensory uncertainty, and outcomes under each policy can determine individual protective behaviours. We suggest that these mechanisms are crucial to assess how individuals cope with uncertainty during a pandemic, and we show the interest of these new perspectives for public health policies.

Keywords Active inference · Bayesian inference · Coronavirus · Protection motivation theory · Health belief model · Pandemic

Introduction

Emerging infectious diseases, such as COVID-19, create new challenges for public healthcare systems. Without a treatment, countering the spread of these diseases depends largely on protective behaviours on the part of individuals and groups, such as social distancing, respecting quarantine, lockdown, wearing masks, frequent handwashing, travel restrictions, and vaccine

acceptance (Steven et al., 2020; Mizumoto & Chowell, 2020; Kretzschmar et al., 2020; Kissler et al., 2020; Friston et al., 2020; Aleta et al., 2020). These individual and social measures can reduce transmission rates, and subsequently alter mortality rates and the number of active cases (Jefferson et al., 2009; Hayward et al., 2020). Mathematical modelling and analysis of individual behaviour at the population scale during previous epidemics suggests that the degree to which protective behaviours are enacted, especially social distancing, effectively predicts the timing, and course of global disease trajectories (Brockmann & Helbing, 2013). Conversely, the inferred prevalence of the virus and accompanying fear predicts behaviour on a population-level (Fast et al., 2015). This straightforward observation has a profound implication; namely, that of a circular causality, in which protective behaviours modulate transmission and spread of infection, while the prevalence of infection in turn induces protective behaviours (Friston et al., 2020; Kassa & Ouhinou, 2015; Cherif et al., 2016).

Previous work suggests that committing to protective behaviours depends on perceived risks (i.e., beliefs about individual

✉ Hugo Bottemanne
hugo.bottemanne@gmail.com

¹ Institut du Cerveau - Paris Brain Institute (ICM), UMR 7225/UMR_S 1127, Sorbonne University/CNRS/INSERM, Paris, France

² Department of Psychiatry, Pitié-Salpêtrière Hospital, Assistance Publique - Hôpitaux de Paris (AP-HP), Paris, France

³ Sorbonne University, Department of Philosophy, SND Research Unit, UMR 8011/CNRS, Paris, France

⁴ Wellcome Trust Centre for Human Neuroimaging, Institute of Neurology, University College London, London, United Kingdom

vulnerability and threat severity) and on the estimated availability and efficacy of protective actions (i.e., beliefs about the efficacy of the response and about people's ability to engage in such protective actions) (Webster et al., 2020; Williams et al., 2015; Teasdale et al., 2014). These models include situations where individuals believe protective behaviours are effective but are not able to enact them (e.g., wanting a mask but not having one, or wanting to social distance but being stuck in prison). They assume that if a person believes that the infection can be prevented, that infection can be avoided with protective behaviour, and that they have the means to perform those behaviours, then this person is more likely to perform protective behaviours (Xiao et al., 2016; Sesagiri Raamkumar et al., 2020; Hu et al., 2017). During a pandemic, information from several sources, such as media, government, and interpersonal relationships, can increase awareness of the risks associated with the disease and whether preventive measures should be adopted (Lin et al., 2020). However, little is known about the genesis of these beliefs in responses to an epidemic and about the cognitive mechanisms that may link these beliefs to decision-making and action. In order to develop sanitary strategies, it may be useful to gain better understanding of the mechanisms that define the association between perceived risk and protective behaviour.

The *active inference framework* is a novel approach to behavioural modelling that integrates embodied perception and action, belief updating, and decision making (Friston et al., 2016; Friston, 2020). It addresses the problem of inferring (unobserved or hidden) states of the world, learning the statistical structure of the world, and acting in an appropriate manner based on a set of preferred outcomes and probabilistic beliefs about an uncertain and changing environment (Friston et al., 2016). Active inference casts action, perception, and cognition as minimizing quantities called variational and expected free-energy. The former minimizes the divergence between predicted and observed sensory outcomes and the latter minimizes the divergence between preferred and observed sensory outcomes of actions—based on a model of how sensory data are generated under distinct plausible behaviours (Friston et al., 2006). According to active inference, the brain deploys a form of (variational or Bayesian) inference to infer the unobserved causes of its sensory data, and to select action sequences (policies) that actively change the world to bring about expected or preferred sensory outcomes (Friston et al., 2006). In other words, action and perception work hand-in-hand to minimise free-energy or to get the agent as close as possible to its preferred sensations.

Active inference provides a formal framework to model how agents update their probabilistic models of the world by collecting sensory data that are generated by the consequences of action. It thereby provides a mathematical description of the cognitive and behavioural adaptation of a biological agent to its environment (Friston, 2013; Ramstead et al., 2018; Constant et al., 2018), and a framework for studying the behaviour of individuals and groups in situations that require

decision making under uncertainty (Friston et al., 2016). This approach is thus valuable for understanding how beliefs determine individual and collective behavior during an unexpected phenomenon such as a pandemic. In this paper, we present a computational account from the AI principles describing how individuals update their beliefs about the risks and thereby commit to protective behaviours. We show how perceived risks, beliefs about future states, sensory uncertainty, and outcomes under each policy can determine individual protective behaviours. We suggest that these mechanisms are crucial to assess how individuals cope with uncertainty during a pandemic, and we describe the interest of these new perspectives for public health policies.

Active inference in the brain?

In the cognitive neurosciences, active inference provides a formal framework for understanding the choice behaviour of individuals under uncertainty. This theory proposes that the dynamics of the brain minimize variational and expected free-energy (Friston et al., 2006). In information theory, free-energy provides an upper bound on self-information (a formal measure of surprise) where expected surprise is known as entropy or uncertainty (see Technical terms for a detailed explanation of terms). Crucially, minimizing free-energy is equivalent to maximizing Bayesian model evidence, i.e., the probability of sensory exchanges with the environment under a model of how those sensations are caused (Friston, 2010). Accordingly, the brain maintains an internal representation of all the relevant statistical variables in the environment. This representation rests on a probability distribution over hidden states—and the observable consequences generated by those variables; such (probabilistic) models are called *generative models* because they represent the causal factors that generate sensory data. Based on sensory observations, the brain can update its representations, with an algorithmic process equivalent to an approximate form of Bayesian inference about the (hidden) state of its environment (Friston et al., 2017). This inference corresponds to minimising free energy or maximising the evidence for the generative model (see Fig. 1 for an example of a generative model in AI).

Active inference assumes that perception and action are two major ways in which free-energy is minimized (Ramstead et al., 2019). Heuristically, perception makes internal representations more like the data that the brain acquires. Reciprocally, action makes the data closer to the preferred distribution that is represented internally. Action (or policy selection) involves inference premised on a generative model that represents the expected sensory consequences of action, where the sequence of actions that is selected is the one that best reduces (expected) free-energy. This is sometimes referred to as *planning as inference* (Botvinick & Toussaint,

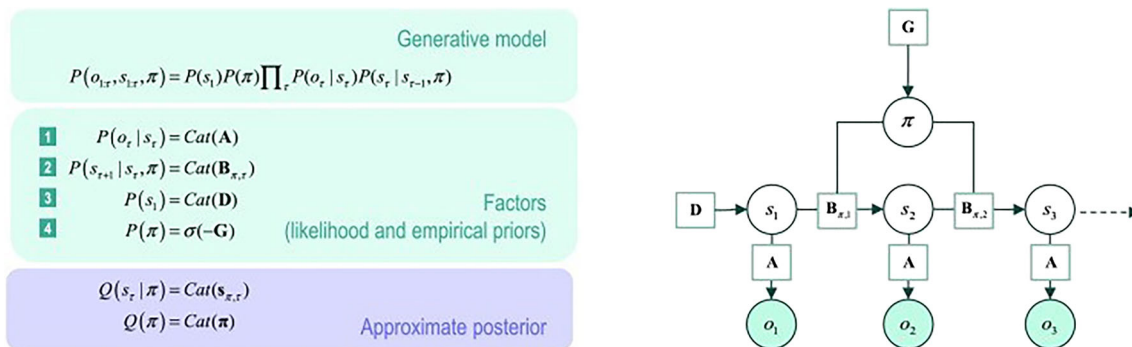


Fig. 1 Example of a generative model in AI (Active Inference). A generative model is a probabilistic specification of how outcomes are caused. Usually, the model is expressed in terms of a likelihood (the probability of consequences given causes) and priors over causes. Examples of these probability distributions are provided in the green boxes. Bayesian model inversion refers to the inverse mapping from consequences to causes; i.e. estimating the hidden states that cause outcomes. In approximate Bayesian inference, one specifies the form of an approximate posterior distribution (blue box) with a specified functional form—that is chosen to make model inversion tractable. Left panel: these equations (in the green boxes) specify the generative model: the likelihood is specified by a matrix A. The elements of A encode the probability of each outcome for each hidden state. Cat means a categorical probability distribution. The priors include probabilistic transitions (in B matrices) among hidden states that can depend upon

actions, which are determined by policies (i.e., sequences of actions denoted by π). The key aspect of this generative model is that policies are more probable a priori if they minimize the expected free energy G, which depends upon prior preferences about outcomes or costs (encoded by C). Finally, the vector D specifies prior beliefs about the initial state. This completes the specification of the model in terms of parameters that constitute A, B, C, and D. Right panel: the accompanying generative model shown as a Bayesian dependency graph: this Bayesian graph depicts the conditional dependencies among hidden states and how they cause outcomes. Open circles are random variables (hidden states and policies), while filled circles denote observable outcomes. Squares indicate fixed or known variables, such as the model parameters (See Friston, Parr, De Vries (2017) for a detailed explanation of the variables and mathematics).

2012). This notion supposes that the brain is capable of predicting the perceptual feedback that would be produced by adaptive actions. It infers the hidden states that cause sensations, and chooses actions that minimize expected surprise in the future (Friston, 2010). For that purpose, the brain accumulates sensory evidence; and perception corresponds to updating probabilistic beliefs or representations about the current state of the world. In action, rather than inferring the causes of sensory data, the brain infers actions that are expected make sensory data accord with its preferences about sensory input (i.e., avoid surprises expected in the future). The value of each policy is then evaluated in terms of its expected free-energy (i.e., surprise), such that the policy that leads to the least expected free-energy is the one that is selected. Behaviour therefore depends on beliefs about future states and outcomes under each policy. Actions realize these predicted outcomes, eliciting new evidence from the world (Friston et al., 2010).

Free-energy minimization has been proposed as an explanation for collective behaviours premised on shared cultural, social, and trans-personal conventions (Ramstead et al., 2016; Constant et al., 2019). Crucially, these accounts explain collective (multiagent or ensemble) behaviour in terms of individual actions premised on a shared generative model. Recall that a generative model specifies the manner in which typical sensory data are caused, especially by action. To share a generative model means to share such sets of expectations. In this way, social conformity comes from individual inferences premised on a shared model and from the enactment of those expectations via environment-

modifying actions (Ramstead et al., 2016; Constant et al., 2019). This computational approach is thus able to describe the way in which individual beliefs determine individual behavior, jointly using knowledge from neuroscience and social psychology. But it is also able to describe how individual beliefs determine collective behavior, and gradually form collective beliefs. During a pandemic, these perspectives thus offer the possibility of better understanding the interaction between the information transmitted within the population, the beliefs of individuals about the pandemic, the uncertainty caused by this unexpected phenomenon, and the actions taken to protect themselves against the virus (aka, the protective behaviours).

How to protect yourself during an outbreak?

Pandemics are massive generators of uncertainty. Infectious diseases are generally perceived as less controllable than chronic diseases, such as diabetes, cancer, or heart disease (Kasperson et al., 1988). When facing an epidemic, individuals appraise the characteristics of the threat itself, and their ability to act against that threat (Niels, 2020). First, the threat is generally assessed in terms of dangerousness. Second, the predicted effectiveness of protective strategies and the perceived vulnerability to infection each modulate the salience of the threat. Proportional to threat salience, one of the central emotional responses to a pandemic is fear (Person et al., 2004; Chang et al., 2004; Depoux et al., 2020; Leung et al., 2005). In

the context of pandemic, people generate a number of beliefs about the risk of infection given large amounts of available, but ambiguous, evidence (Lau et al., 2006; Lau et al., 2007). These evidence come from local sources that are communicated through personal connections, and global sources that depend on extrinsic factors like the media (Xiao et al., 2016; Sesagiri Raamkumar et al., 2020; Funk et al., 2010). According to these sources, people update beliefs about the chances of being infected, the seriousness of the disease, the efficacy and availability of protective actions, and their ability to commit to such protective actions (Rogers, 1975; Lau et al., 2003; de Zwart et al., 2009). Numerous studies have shown that these beliefs are a major determinant of protective behaviours, such as social distancing, mask wearing, and handwashing, or for the respect afforded to collective rules, such as mandatory quarantine, lockdown, and travel restrictions (Fast et al., 2015; Kassa & Ouhinou, 2015; Cherif et al., 2016; Webster et al., 2020; Teasdale et al., 2014).

The brain continuously mobilizes beliefs about the severity of the illness, the probability of infection, the efficacy of the behaviour to reduce the probability of infection, and finally, the probability of infection if a new behaviour is adopted. Individuals could reduce their fear by updating these beliefs. This conception may be associated with previous models of perceived risk that distinguish between an “automatic” emotional reaction (a quick and automatic feeling about risk), and a slower cognitive reaction (a more explicit, calculative appraisal of risk) (Slovic et al., 2004; Dillard et al., 2018). The first automatic type of response corresponds to belief updating about the risk, i.e., increased estimated likelihood of risk; whereas the second, deliberative response may be more associated with explicit policies (beliefs about actions), i.e., appraisal of coping strategies. These reactions (updating beliefs vs. action-oriented decision making) could then be qualified as adaptive (e.g., following group advice, seeking information) or maladaptive behaviours (e.g., denial of risk, avoiding new information).

Active inference offers an attractive framework for integrating uncertainty, emotion, belief, and action (Friston, 2010). Interestingly, most theories based on active inference associate negative emotions, such as fear, with inferences about increases in (expected) free-energy over time, where expected free energy (i.e., surprise) can be read as uncertainty (Hesp et al., 2020). In this computational formulation of *affective* inference, a hierarchical generative model is used in which negative affective states are modelled as “states of self,” i.e., higher-order states that are inferred on the basis of lower-order beliefs (i.e., “I must be stressed because I can't decide what to do next”). The expected free-energy can then be read as a kind of internal estimate of “how well I am doing”—such that increases in expected free-

energy suggests poor performance, i.e., a failure to resolve uncertainty or realise preferred sensory states.

Crucially, an agent can evaluate the degree to which it trusts the expected free-energy that it is generating. Heuristically, if consistently higher-than-anticipated levels of expected free-energy are generated under that model, then it is not a particularly good model. Under this conception, negative emotional states indicate that the agent's attempts to secure preferred outcomes have been consistently thwarted. This entails that the agent's predictive grip on its world is lacking. Negative emotion is therefore a sign that the agent is losing its predictive grip. This model of emotional valence—as hierarchical inference about irreducible expected free-energy (i.e., uncertainty)—provides an account of how emotional states nuance posterior beliefs during Bayesian inference (Hesp et al., 2020). This is usually cast in terms of emotional states predicting the predictability or *precision* of lower level representations.

A loss of certainty about states of affairs—and how to respond—corresponds to a loss of precision in representations or probabilistic beliefs. Precision is an important attribute of probabilistic beliefs and can be thought of as the opposite of uncertainty (e.g., inverse variance). In active inference, the precision of a belief has itself to be inferred, where this inference corresponds to attention. In other words, affording a representation greater precision corresponds to attentional selection. Technically, the precision assigned to various sources of evidence is a key quantity in Bayesian inference, ensuring that more reliable sources of evidence contribute to belief updating.

Beliefs about coping responses are crucial for understanding the mechanisms involved in emotional responses to pandemics (Boer & Seydel, 1996). Beliefs about threat depend on evaluating the state of the environment and observing what happens to individuals, whereas beliefs about coping are compelled by the perceived response efficacy (the belief that the recommended behaviour will be protective) and one's own self-efficacy (the ability to perform the recommended behaviour). Individuals evaluate whether a protective action will mitigate the threat (response efficacy), their level of confidence in being able to carry that action out (self-efficacy), and also the cost of this protective action. Beliefs about coping responses encompass beliefs about the threat, because if an individual is convinced that they could protect themselves from risks, this reduces fear (and uncertainty) associated with the threat. Active inference allows us to integrate beliefs about the risk (threat appraisal), beliefs about action (coping appraisal), and decision making. It suggests that the brain could minimize free-energy (i.e., beliefs about high vulnerability and beliefs about high severity) by fulfilling

its prediction about the availability of protective actions, and predictions about the ability to engage in such protective actions.

Towards an active inference account of protective behaviour

This model of active inference is consistent with previous theories associating health beliefs and protective behaviours. Traditionally, two competing theories of health-protective behaviour are proposed to explain the link between beliefs and health actions: the health belief model (HBM) and the protection motivation theory (PMT) (Rogers, 1975; Janz & Becker, 1984). They are value-expectancy theories, based on the assumptions that people want to avoid illness and believe that behaviours will prevent illness. They describe the cognitive processes that mediate behaviour in the face of a threat and suppose that the motivation to protect oneself is the proximal determinant of these behaviours (Maddux & Rogers, 1983). These theories have been applied extensively in numerous frameworks in medicine, nutrition, or cybersecurity, to predict various health-related behaviours (Milne et al., 2000). The health belief model (HBM) assumes that health behaviour depends on the perceived threat, the perceived benefits of health behaviour, and the perceived cost of health behaviour, comprising economic, social, and psychological costs, such as anxiety, shame, or discomfort (Rosenstock, 1974; Becker et al., 1977). Protection motivation theory (PMT) supports HBM theory by incorporating several additional factors. It refines these assumptions and includes beliefs about self-efficacy as well as the conviction that one can successfully execute the behaviour required to produce predicted outcomes (Rogers, 1975; Bandura, 1997).

PMT suggests that people's motivation to engage in precautionary actions is influenced by two major factors: threat appraisal and coping appraisal (Rogers, 1983). Threat appraisal encompasses beliefs about vulnerability, a subjective estimate of the chances of contracting a disease (how likely one is to get the illness), and beliefs about the severity of a disease (how serious the illness is). Coping appraisal involves beliefs about the efficacy of the response, protective actions that are available, and beliefs about self-efficacy; that is, one's ability to effectively engage in such protective actions (Rogers, 1975). PMT thus includes the beliefs of individuals regarding their ability to respond to the threat, and about the efficacy of this response to that threat (Maddux & Rogers, 1983). Explicitly, these beliefs cover the self-vulnerability ("my chances of getting Covid-19 are high"), the severity of the pathogen ("pneumonia resulting from coronavirus is a serious condition"), the perceived benefits ("handwashing could reduce the risk of contracting COVID-19"), but also the cost

involved in carrying out that behaviour ("social distancing will make me sad").

These beliefs can be represented as a hierarchy like the one we proposed in our model of active inference. In this hierarchy, each belief is embedded in a network of causal dependencies: beliefs about the ability of individuals to protect themselves influence beliefs about individual vulnerability or about the dangerousness of a pathogen. If you believe that potential infection will only have mild effects on you, then the perceived efficacy of coping strategies is reduced (because they reduce an already low risk). If you believe that coping strategies are effective, then your perceived potential risk of infection must be low. This dependence ensures that one's motivation to remain healthy can influence coping appraisal, predicted vulnerability, and perceived severity, but also perceived benefits and predicted costs. In other words, the motivation to realize protective actions may sensitize individuals to threat signals, while a lack of motivation may desensitize individuals to such signals. On the other hand, the consequences of action could influence the strength of protective motivations. If, after realizing action, perceived costs were higher than perceived benefits, the perceived threat would be expected to increase. In turn, the perceived threat could strengthen motivation to enact protective behaviours, leading to a self-reinforcing feedback loop.

These theories have been successfully applied to understand the evolution of behaviour during past outbreaks of severe acute respiratory syndrome coronavirus (SARS) (Jiang et al., 2009; Tang & Wong, 2005; Tang & Wong, 2003), influenza A virus subtype H1N1 (H1N1) (Cowling et al., 2010; Sharifirad et al., 2014; de Zwart et al., 2010), and Middle East respiratory syndrome (MERS-CoV) (Alqahtani et al., 2017; Jang et al., 2020). In the 2009 H1N1 outbreak, the belief that the virus could be spread by indirect contact was directly associated with a greater use of hygienic measures and social distancing (Cowling et al., 2010). This belief increased the perceived vulnerability of an individual, but also indirectly increased the predicted response efficacy and self-efficacy. During the early stages of the COVID-2019 pandemic, a study conducted in South Korea showed that the enactment of precautionary behaviours was strongly associated with perceived risks, and with beliefs about the efficacy of those behaviour (Lee & You, 2020). The majority of respondents (51.3%) reported that their perceived risk of infection was "neither high nor low"; 48.6% reported that they believed that the severity of illness would be "high"; 19.9% reported that they believed it would be "very high." In this sample, 41.5% were avoiding crowded places, 50% reported cancelling social events, 63.2% reported always wearing a facial mask, and 67.8% reported always practicing hand washing. Interestingly, the average perceived severity score was higher than perceived vulnerability. Another study that was conducted in Iran showed that both threat appraisal and coping

appraisal predicted protective behaviour (Barati et al., 2020). These findings suggest that, whether for old epidemics or for the current coronavirus crisis, a high perceived risk and a good coping appraisal are strongly associated with social response.

Adaptive and maladaptive behaviours as free-energy minimization

In our active inference model, optimal behaviour entails actions that resolve uncertainty and achieve preferred, unsurprising outcomes. These behaviours can be specified in terms of policies that minimize the free-energy expected when pursuing them (Friston et al., 2016; Friston et al., 2010). If the brain seeks to minimize free-energy, it can select protective behaviours, which fulfil predictions about the threat and about coping responses. In short, the brain chooses policies that minimize uncertainty about future outcomes, by minimizing the free-energy expected following action. Crucially, expected free-energy can be decomposed into epistemic and pragmatic terms that can be alternatively expressed in terms of *risk* and *ambiguity*.

Here, risk scores the difference between predicted and preferred outcomes, where preferred outcomes and those that are least surprising a priori (e.g., avoiding infection). Ambiguity reflects the uncertainty about observations, given their causes. Therefore, choosing policies to minimise expected free-energy maximizes preferred outcomes while, at the same time, avoiding ambiguous situations, such as “being in the dark” (i.e., this has the effect of driving the agent to seek the most salient or informative observations). These two aspects of expected free-energy can be regarded as uncertainty of a specific and nonspecific sort that pertain to specific prior expectations about preferences (i.e., risk) and a generic ability to infer states of the world (i.e., ambiguity). Mathematically, risk corresponds to the expected “cost” of a policy.

On this account, the outcomes of protective behaviours confirm beliefs about the ability of coping to achieve preferred outcomes and reduce uncertainty. Reduction of negative affect (e.g., fear) via free-energy minimization will reinforce protective behaviour in the future (Friston et al., 2010). The realization of protective behaviours thus forms a loop of belief-confirmation and epistemic habits. This account could explain why protracted experience of threat reduces perceived risk, in the sense that fulfilment of predictions about coping strategies reduces uncertainty about policies, independently of some variation in the threat itself (Savadori et al., 1998). Protection motivation arises when beliefs about response efficacy and self-efficacy outweigh cost, and when protective actions effectively fulfil predictions about severity and vulnerability. Dovetailing with this account, studies performed during SARS epidemics show that the perceived response efficacy and self-efficacy were strongly associated with protection

behaviours (Jiang et al., 2009; Tang & Wong, 2005; Tang & Wong, 2003). During the COVID-19 pandemic, results from a national cross-sectional online survey of 1420 Australian adults in March 2020 show that protection behaviours were associated with a higher rating of perceived effectiveness of behaviours and higher levels of perceived ability to adopt social distancing strategies (Seale et al., 2020). This association between precautionary behaviours, perceived severity and perceived self-efficacy was found in a study that investigates the impact of online information during the pandemic (Farooq et al., 2020).

However, to minimize free-energy, the brain can adopt other strategies, especially when there are no plausible coping policies. These strategies encompass maladaptive behaviour, such as avoidance, or maladaptive beliefs, such as denial or wishful thinking (Conner & Norman, 2005; Witte & Allen, 2000). In these cases, individuals entertain internal actions or policies (e.g., attentional mechanisms) that control emotions rather than limiting risks—which also minimizes free-energy, albeit maladaptively if the resulting beliefs are not attuned to the real risks present in the environment (Loewenstein et al., 2001). In other words, by changing the confidence or precision afforded certain beliefs, the brain can effectively ignore sources of evidence, leading to a decrease in the perceived threat (e.g., ignoring cues that would otherwise suggest the situation is dangerous). These maladaptive behaviours and beliefs allow together individuals to avoid or reduce the threat, and therefore maintain a low free-energy. High perceived risk will elicit protective behaviour only when the individual has sufficient confidence about coping efficacy. If there is high uncertainty about these coping strategies, perceived risk may produce a greater level of maladaptive responses.

In the early phase of the outbreak, the discrepancy between predictions and sensed outcomes will largely increase free-energy, producing uncertainty and fear (Liao et al., 2014; Courtney et al., 2020). Faced with this uncertainty, the choices of policies are limited. The need to reduce uncertainty may encourage individuals to ignore (i.e., reduce the precision of) evidence of risk, resulting in some reassuring underestimation of the severity of the epidemic (Dolinski et al., 2020). These attenuated or biased beliefs can be understood as nuancing the risk (and ambiguity) of the world. In effect, this kind of—possibly some personal—denial is a Bayes-optimal response to a world that cannot be predicted or explained. This phenomenon could explain the discrepancy between sensory evidence and people’s cognitive representations of risk at the beginning of COVID-19 epidemic (Savadori et al., 1998; Loewenstein et al., 2001; Raude et al., 2020). In February and the beginning of March, a large proportion of the European public did not consider the novel coronavirus to be a significant threat—sometimes attributed to an unwarranted exceptionalism (i.e., “This could not possibly happen to us”) (Betsch, 2020). This collective denial has been difficult

to understand, given the accelerating death count in China, Italy, and France (Balogun, 2020). During this period, many Europeans developed maladaptive beliefs, such as “the virus is like influenza,” “it only affects old people,” or “it will never come through the border.” In the same way, while epidemiologists pointed out risks of infection ranging from 11% to 19% during the 2009 H1N1 outbreak, the majority of people believed that they were unlikely to get infected and to infect others (Xu & Peng, 2015), and felt that the pandemic did not affect their daily habits (Lau et al., 2009).

This active inference principle may give some insights about the production of conspiracy beliefs during COVID-19 pandemic (such as denial of the existence of the virus or denial of its dangerousness) and the ability to adopt preventive and protective behaviours. In front of high uncertainty, individuals can accumulate information transmitted by peers, medias, and social networks about the pandemic, but their action will remain limited to these information channels. They will not be able to go to Wuhan to investigate the site of the first contaminations, question scientists to hear their conclusions, or even find scientific evidence about this origin on their own: carrying out actions (internet research, reading the press, conversation with peers) remains associated with a high degree of uncertainty. Alternatively, conspiracy beliefs propose an explanation of these ambiguous phenomena for individual and may participate to reduce this uncertainty. Studies showed that conspiracy theorists are less likely to adopt normative preventive behaviours when they perceive a significant risk for themselves (Allington et al., 2020; Earnshaw et al., 2020; Imhoff & Lamberty, 2020). This effect could be linked to the lack of need for these subjects to reduce uncertainty by carrying out actions: uncertainty is already sufficiently minimized by conspiracy beliefs, and protective actions become useless. Linked to this hypothesis, studies also suggest that individuals who present conspiratorial beliefs are less likely to adhere to vaccination (Morgan & Laland, 2012). This hypothesis explains why conspiracy beliefs not involving the dangerousness of the virus (e.g., conspiracy beliefs about the origin of the virus) restrain the observance of protective measures. These conspiracy beliefs could minimize uncertainty, reducing the need for action. In the competition between generating beliefs to better understand the world, and taking action to acquire more precise information about the world, conspiracy beliefs thus fulfill their role.

Limits and theoretical perspectives

The computational hypotheses that we have detailed deserve to be compared with alternative theories concerning protective behaviours. As we described earlier, the two most used theories in social psychology and public health are health belief model (HBM) and protection motivation theory (PMT)

(Rogers, 1975; Janz & Becker, 1984). These models assume that health behaviour depends on threat and coping appraisals, encompassing beliefs about vulnerability, severity, efficacy of the response, and self-efficacy (Cowling et al., 2010; Sharifirad et al., 2014; de Zwart et al., 2010). In this paper, we have shown how the principles of active inference may enlighten the HBM and PMT mechanisms and bring new ways of understanding experimental results in social psychology.

Our theory thus links the perceived level of the threat to the confidence (i.e., precision afforded) in the efficacy of coping strategies. If the level of the threat is believed to be too high, and the precision of expected outcomes from protective policies is believed to be too low, the threat itself could then inhibit the protective action. This phenomenon is particularly apparent in the context of screenings for serious diseases, such as cancer or HIV (McCaskill, 2006). Screening tools then provide information on the risk, but if the risk is confirmed, the individual knows that the possibility of reducing risk is low. In other words, there is a balance between reducing the uncertainty of risk and the risk of increasing uncertainty. Avoidance and denial are then a quick and effective way to resolve uncertainty. By not performing the screening test, the individual protects himself from the possibility of bad news. The fear of death and physical pain—as a result of being diagnosed with cancer or HIV—can lead to denial and hinder screening behaviours.

Otherwise, unlike HBM and PMT models which assume that attitudes (and actions) are determined by the conjunction of evidence supporting behaviour and beliefs about the usefulness of those behaviours, our theory emphasizes the primacy of action and uncertainty minimization. This account agrees with studies suggesting that individuals first develop attitudes and then find evidence to support those attitudes (Dong 2008). In this setting, an individual wears a protective mask if this action minimizes uncertainty, and beliefs about the ability of the mask to protect against the virus (or the usefulness of protection against the virus) are directly associated with this tendency to minimize uncertainty. This mechanism coming from action to beliefs also may protect beliefs that are incompatible with acquired sensory evidence and emphasizes the primacy of action over beliefs generation.

Moreover, extensions of active inference to social phenomena may shed light onto individual reactions to the uncertainty of the pandemic. Facing a pandemic, individuals coordinate and cooperate with each other, supporting group decisions based on shared goals. Beliefs about risks and protective actions are a crucial part of the shared expectations—to which an individual or group implicitly appeals when they choose a behavioural policy. Individuals acquire these expectations through shared experiences in a social or epistemic community; individual behaviour is mainly shaped by the social subgroups which embrace their social identities (Turner et al.,

1987). These shared expectations could explain social conformity in protective behaviours, by generating automatic behavioural responses in accordance with preferences, values, and goals that are features of an epistemic community (Van Assche et al., 2020). For the group, the relevance of protective actions therefore may have been associated with beliefs about generalised individuals vulnerability, and in return, the display of protective actions in the group may sensitize individuals to threat signals.

However, others theoretical frameworks in cognitive sciences may offer some understandings of the emergence of adaptive and maladaptive behaviours toward health threats (Bavel et al., 2020; Bonell et al., 2020). For example, psychological theories focusing on individual psychological characteristics are particularly interesting for explaining protective behaviour in the population. Studies have shown that individuals with reduced ability to think about the future are more likely to engage in healthy behaviour when positive outcomes are immediate, and negative outcomes are seen as only having effects on the long term (Orbell et al., 2004). Alternatively, individuals with a greater ability to think about and project themselves into the future are more likely to undertake health behaviours when immediate outcomes are negative, and long-term outcomes are positive (Orbell et al., 2004). These results can explain behaviours during the pandemic and are difficult to reconcile with the principles of active inference.

Otherwise, studies have shown that compliance to protective behaviour result from people's capacity to obey the rules, opportunity to break rules, and people's intrinsic motivations, comprising moral support and social norms (Wolf et al., 2020; Yamamoto et al., 2021). These individual psychological characteristics are only partially taken into account in the principles of active inference. For example, individuals share group membership with other agents during the pandemic: in this setting, protective behaviours may become a socially approved norm, and social conformity is an adaptive strategy to cope with this kind of uncertain environmental situation (Morgan & Laland, 2012; Cruwys et al., 2020). This cognitive mechanism includes the tendency towards increased appraisal of information from socially relevant agents (e.g., people who elicit epistemic trust) and to imitate these agents (Laland, 2018; Dong, 2008). Moreover, some altruistic behaviour are observed for individuals who have low personal risk (e.g., younger age, without preexisting medical conditions) but could still directly or indirectly transmit the virus to high-risk people (Galang et al., 2021; Walker, 2021). These individuals adopt (and maintain) nonpharmaceutical mitigation strategies to protect others, primarily their loved ones, but sometimes also society as a whole (Walker, 2021; Petrocchi et al., 2021). The principles of active inference do not yet allow us to understand these altruistic behaviours facing infectious agents, and these behaviours should be explored by further studies.

Other lines of evidence also suggest that maladaptive behaviours can carry reward, such as intrinsic pleasure and social approval, which can contribute to maintain maladaptive behaviour (Conner & Norman, 2005). This account is crucial to understand the factors that determine whether people accept or reject control or protection measures (Reicher & Stott, 2020). Indeed, some alternative beliefs (e.g., belief that the virus is safe) or certain nonnormative behaviours (e.g., to decline wearing a mask against the virus) can reinforce the self-esteem of individuals. This rewarding effect has been associated with the individual and group self-esteem studied in psychology. In this setting, generating maladaptive beliefs can give the feeling of holding a truth, and thus increases the feeling of being special (Cichocka et al., 2016; Golec de Zavala & Federico, 2018). This highlights a complicated interplay between prior preferences (that determine risk) and the need for clarity (that resolves ambiguity) when selecting a course of action (Constant et al., 2018; Ramstead et al., 2016; Constant et al., 2019; Veissière et al., 2020).

Moreover, perceived risks and protective behaviours may be linked with other cognitive biases, such as proportionality bias and optimism bias (Bottemanne et al., 2020). Proportionality bias designates the tendency to modify beliefs and behaviours according to the importance of a phenomenon. For example, the increasing severity of a phenomenon, such as Sars-CoV2 infection, is generally associated to the number of beliefs theories about this phenomenon. This bias implies that major events must be explained by major causes and be prevented by major measures. According to this principle, the propensity to adopt protective behaviours may be directly associated with the perceived importance of the pandemic. Optimism bias designates the tendency to update more beliefs after positive information, rather than negative ones (Sharot, 2011; Sharot & Garrett, 2016). As we have seen, a large proportion of individuals did not consider the novel coronavirus to be a significant threat at the beginning of the pandemic (Betsch, 2020), given the accelerating deaths count in China and Europe (Balogun, 2020). This optimism bias has been proposed to explain the discrepancy between sensory evidence and people's cognitive representations of risk at the beginning of COVID-19 epidemic (Savadori et al., 1998; Loewenstein et al., 2001; Raude et al., 2020). However, we have suggested that this cognitive bias may be associated with uncertainty minimization: the limited choices of policies to reduce uncertainty may encourage individuals to ignore (i.e., reduce the precision of) evidence of risk, resulting in a reassuring underestimation of the severity of the pandemic (Dolinski et al., 2020). These optimistic biased beliefs can be understood as shading the risk (and ambiguity) of the world. Taken together, these alternative hypotheses are not fully explain by the principles of active inference and deserve to be considered as alternative cognitive theories for protective behaviours during a pandemic.

From theory to guidelines for global health policies

Despite its limits, active inference may furnish some perspectives on developing guidelines for global health strategies. In response to the COVID-19 pandemic, governments across the world have adopted measures to slow the spread of the virus such as lockdown and vaccine. The behavioural response of individuals during epidemic is one of the major variables to limit the spread and to predict the effects of societal or institutional responses to the epidemic (Friston et al., 2020; Durham & Casman, 2012). Collective measures targeting these behavioural responses are critical to decrease mortality and reduce the overburdening of healthcare systems (Leung et al., 2009). For example, vaccine hesitancy became a major political and health issue, limiting or slowing the protection of populations (Khubchandani & Macias, 2021).

As we have seen, individuals coordinate and cooperate with each other, supporting group decisions based on shared goals regarding the pandemic. Beliefs about risks and protective actions are a crucial part of these shared expectations; individuals acquire these expectations through shared experience in an epistemic community (Turner et al., 1987). These latter could explain social conformity in protection behaviour, by generating behavioural responses in accordance with preferences, values, and goals that are specific to the epistemic community (Van Assche et al., 2020). The relevance of protective actions therefore may have been associated with beliefs about generalised individual vulnerability, and in return, the display of protective actions in the group may sensitize individuals to threat those signals.

Later, protective behaviour may be characterised by environmental cues that denote specific actions to be accomplished, given that certain cues are perceived (Constant et al., 2019). The idea here is that observing our peers increases the likelihood that we will engage in the same behaviour. If this is correct, governments should promote the spread of positive appraisals of public protective behaviour (e.g., the obligation to wear a mask in public places) to produce “epistemic pressure” leading to behavioural conformity (Betsch, 2020; Bavel et al., 2020). In line with this view, a study performed during the spread of COVID-19 revealed that promoting collectivism may be a way to increase engagement in protection behaviours (Biddlestone et al., 2020).

Moreover, if action fulfils predictions based on perceptual inference, and if the brain favours actions that minimize expected free-energy, then the prediction of the effectiveness of the action—in relation to risk and ambiguity—is crucial. Therefore, the major role of health policies and communication should be to increase the precision of various beliefs, i.e., resolve uncertainty about the expected outcomes associated with protective behaviour. Namely, such beliefs have to underwrite confidence about coping outcomes. Accordingly, a

meta-analysis showed that both threat and coping appraisals were significantly associated with protective behaviour, but this association was stronger for coping appraisal than threat appraisal, and especially for self-efficacy (Milne et al., 2000). Alternatively, self-efficacy and response efficacy were negatively correlated with maladaptive behaviour which inhibits protection motivation.

These active inference principles may also give some insights about vaccine hesitancy. Vaccine acceptance and hesitancy depend mainly on beliefs about the risk of the virus, and beliefs about the vaccination including: safety, harmlessness, efficacy of vaccination, but also confidence in the practitioner providing it, in the pharmaceutical laboratory producing it, and more generally in the health system (Xantus et al., 2021). The content of these beliefs (e.g., the low risk of contracting the disease, and the high risk associated with vaccination) contribute to a noncompliance of the government recommendations. The fight against vaccine hesitation can take several ways: first, by limiting the generation of maladaptive beliefs about the virus, which allow the individual to directly reduce uncertainty without the necessity to take action after; second, by limiting the generation of maladaptive beliefs about the vaccines, which may constitute a brake on the initiation of vaccination (action); finally, by amplifying the perception that vaccination can effectively reduce the risk of contamination, which increasing the capacity of vaccination (action) to reduce uncertainty.

In this setting, governments may have to improve their health message about protective measures and vaccination (Finset et al., 2020; Yousuf et al., 2020; Leung et al., 2020; Piltch-Loeb et al., 2017). During COVID-19 outbreak, a study including 9,000 citizens in Italy dramatically shows that mass-media communication plays a major role in updating these beliefs about pandemic (Motta Zanin et al., 2020). It is crucial for public-health policies to deliver mass health advice, not only about risk, but especially about the effectiveness of protective behaviours. The balance between adaptive and maladaptive response toward an infectious disease depends on the balance between the threat appraisal and the expected risk. Then, the predicted cost, both economic and cognitive, has to be presented as low to facilitate these actions. Individual behaviours, such as handwashing, mask wearing, social distancing, and vaccination, *must be framed as effective* (when they are found to be effective).

Conclusions

Cognitive and theoretical neuroscience may have something useful to offer when fighting the COVID-19 outbreak. Pandemics emerge from interactions between pathogens and epistemic agents. Collective and individual protective measures especially require a fundamental shift in human beliefs

and behaviour. Although we are well aware of the biological processes involved in the propagation of most pathogens, it is difficult to model the cognitive and the behavioural processes of individuals. Insights from the computational and social neurosciences are then critical to enrich models of public health intervention strategies.

Active inference offers a unifying framework to understand how the individuals generate beliefs about risks and commit to protective actions. It assumes that the brain minimizes expected free-energy, a proxy for uncertainty. For that, the brain constantly makes inferences to predict the consequences of action, and update its beliefs based on what the senses relay back. With action, the brain actively samples the world to ensure its predictions become a self-fulfilling prophecy. Protective behaviours, but also maladaptive behaviours, may furnish a way to reduce this uncertainty. In this work, we suggest that increasing the perceived efficacy of protective behaviour, i.e., increasing the precision of beliefs about a consequence of protective action, is a priority for our collective fight against COVID-19.

This formulation acts as a bridge between theoretical models of cognition and epidemiological models—and offers a perspective on the importance of beliefs and believing mechanisms in mitigating pandemic. Mathematical formulations of these cognitive mechanisms could improve the predictive validity of computational models used in epidemiology, if suitably equipped with behavioural responses. Our brain possesses a set of prodigious adaptive systems to fight against ecological threats; it is up to us to understand them, so that we can improve our defences.

Technical terms

- Generative model: a probabilistic mapping from causes to observed data. It is typically specified in terms of the likelihood of getting some data given their causes (i.e., hidden states and parameters of a model) and priors over the states and parameters.
- Bayesian model inversion: Bayesian model inversion refers to the inverse mapping from consequences to causes, i.e. estimating the hidden states that cause outcomes. In approximate Bayesian inference, one specifies the form of an approximate posterior distribution to make inversion analytically tractable.
- Free energy: An information theory measure that bounds (is greater than) the surprise on sampling some data, under a generative model. Minimising free energy by changing the approximate posterior provides a simple and biologically plausible way of inverting a generative model (a.k.a. belief updating).
- Prior: The probability distribution over the causes of data that constitute beliefs about those causes prior to

observing the data. The probability distribution over the causes of data that constitute beliefs about those causes prior to observing the data.

- Posterior distribution: the probability distribution of causes (i.e., hidden states and parameters), given some data; that is, a probabilistic mapping from observed data (consequences) to causes.
- Surprise: Surprise (a.k.a., surprisal or self information) is the negative log probability of an outcome. An improbable outcome is therefore surprising. The log probability of an outcome is also known as Bayesian model evidence (a.k.a., marginal likelihood).
- Precision: the inverse variance or dispersion of a random variable. It quantifies the degree of certainty about the variables (i.e., their predictability).

Acknowledgments The authors thank Maxwell Ramstead, Ryan Smith, Casper Hesp, Fabien Vinckier, Julie Tang, and Lucie Joly for their invaluable help.

Funding This work was conducted under funding for the Wellcome Centre for Human Neuroimaging (Ref: 088130/Z/09/Z).

Declarations

Conflict of interest The authors have nothing to disclose related to this work.

References

- Aleta, A., Martín-Corral, D., Pastore, Y., Piontti, A., Ajelli, M., Litvinova, M., Chinazzi, M., Dean, N. E., Halloran, M. E., Longini Jr., I. M., Merler, S., Pentland, A., Vespignani, A., Moro, E., & Moreno, Y. (2020). Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19. *Nature Human Behaviour*, 4(9), 964–971.
- Allington, D., Duffy, B., Wessely, S., Dhavan, N., & Rubin, J. (2020). Health-protective behaviour, social media usage and conspiracy belief during the COVID-19 public health emergency. *Psychological Medicine*, 9, 1–7.
- Alqahtani, A. S., Rashid, H., Basyouni, M. H., Alhawassi, T. M., BinDhim, N. F. (2017). Public response to MERS-CoV in the Middle East: iPhone survey in six countries. *Journal of Infection and Public Health*, 10(5).
- Balogun, J. A. (2020). Lessons from the USA Delayed Response to the COVID-19 Pandemic. *African Journal of Reproductive Health*, 24(1).
- Bandura, A. (1997). *Self-efficacy: The exercise of Control*. W.H. Freeman.
- Barati, M., Bashirian, S., & Jenabi, E. (2020). Factors Associated with Preventive Behaviours of COVID-19 among Hospital Staff in Iran in 2020: An Application of the Protection Motivation Theory. *The Journal of Hospital Infection*, 0195-6701(20), 30210–30213.
- Bavel, J. J. V., Baicker, K., & Boggio, P. S. (2020). Using social and behavioural science to support COVID-19 pandemic response. *Nature Human Behaviour*, 4(5).

- Becker, M. H., Maiman, L. A., & Kirscht, J. P. (1977). The Health Belief Model and prediction of dietary compliance: A field experiment. *Journal of Health and Social Behavior*, 18(4), 348–366.
- Betsch, C. (2020). How behavioural science data helps mitigate the COVID-19 crisis. *Nature Human Behaviour*, 4(5).
- Bottemanne H, Morlaàs O, Schmidt L, & Fossati P. (2020). Coronavirus: Predictive brain and terror management. *Encephale*, 46(3S), S107–S113
- Biddlestone, M., Green, R., & Douglas, K. M. (2020). Cultural orientation, power, belief in conspiracy theories, and intentions to reduce the spread of COVID-19. *The British Journal of Social Psychology*, 59(3).
- Boer, H., & Seydel, E. R. (1996). Protection motivation theory. In M. Connor & P. Norman (Eds.), *Predicting Health Behavior*. Open University Press.
- Bonell, C., Michie, S., & Reicher, S. (2020). Harnessing behavioural science in public health campaigns to maintain “social distancing” in response to the COVID-19 pandemic: key principles. *Journal of Epidemiology and Community Health*, 74(8).
- Bottemanne, H., Morlaàs, O., Fossati, P., & Schmidt, L. (2020). Does the Coronavirus Epidemic Take Advantage of Human Optimism Bias? *Frontiers in Psychology*, 11, 2001.
- Botvinick, M., & Toussaint, M. (2012). Planning as inference. *Trends in Cognitive Sciences*, 16(10), 485–488.
- Brockmann, D., & Helbing, D. (2013). The hidden geometry of complex, network-driven contagion phenomena. *Science*, 342, 1337–1342.
- Chang, H. J., Huang, N., & Lee, C. H. (2004). The impact of the SARS epidemic on the utilization of medical services: SARS and the fear of SARS. *American Journal of Public Health*, 94(4).
- Cherif, A., Barley, K., & Hurtado, M. (2016). Homo-psychologicus: Reactionary behavioural aspects of epidemics. *Epidemics*, 14, 45–53.
- Cichocka, A., & Marchlewska, M. (2016). Golec de Zavala A. Does self-love or self-hate predict conspiracy beliefs? Narcissism, self-esteem, and the endorsement of conspiracy theories. *Social Psychological and Personality Science*, 7(2), 157–166.
- Conner, M., & Norman, P. (2005). *Predicting Health Behaviour: Research and Practice with Social Cognition Models* (2nd ed.). Open University Press.
- Constant, A., Ramstead, M. J. D., & Veissiere, S. P. L. (2018). A variational approach to niche construction. *Journal of the Royal Society Interface*, 15(141).
- Constant, A., Ramstead, M. J., Veissière, S. P., & Friston, K. J. (2019). Regimes of expectations: an active inference model of social conformity and human decision making. *Frontiers in Psychology*, 10, 679.
- Courtney, E. P., Goldenberg, J. L., & Boyd, P. (2020). The contagion of mortality: A terror management health model for pandemics. *The British Journal of Social Psychology*, 59(3).
- Cowling, B. J., Ng, D. M., & Ip, D. K. (2010). Community psychological and behavioral responses through the first wave of the 2009 influenza A(H1N1) pandemic in Hong Kong. *The Journal of Infectious Diseases*, 202(6).
- Cruwys, T., Stevens, M., & Greenaway, K. H. (2020). A social identity perspective on COVID-19: Health risk is affected by shared group membership. *The British Journal of Social Psychology*, 59(3).
- de Zwart, O., Veldhuijzen, I. K., Elam, G., Aro, A. R., Abraham, T., Bishop, G. D., et al. (2009). Perceived threat, risk perception, and efficacy beliefs related to SARS and other (emerging) infectious diseases: results of an international survey. *International Journal of Behavioral Medicine*, 16(1), 30–40.
- de Zwart, O., Veldhuijzen, I. K., Richardus, J. H., & Brug, J. (2010). Monitoring of risk perceptions and correlates of precautionary behaviour related to human avian influenza during 2006 - 2007 in the Netherlands: results of seven consecutive surveys. *BMC Infectious Diseases*, 10, 114.
- Depoux, A., Martin, S., & Karafillakis, E. (2020). The pandemic of social media panic travels faster than the COVID-19 outbreak. *Journal of Travel Medicine*, 27(3).
- Dillard, J. P., Yang, C., & Li, R. (2018). Self-regulation of emotional responses to Zika: Spiral of fear. *PLoS One*, 13(7).
- Dolinski, D., Dolinska, B., & Zmazyńska-Witek, B. (2020). Unrealistic optimism in the time of coronavirus pandemic: may it help to kill, if so-whom: disease or the person? *Journal of Clinical Medicine*, 9(5).
- Dong, W. (2008). Beyond SARS: ethnic community organization's role in public health – a Toronto experience. *Promotion & Education*, 15(4).
- Durham, D. P., & Casman, E. A. (2012). Incorporating individual health-protective decisions into disease transmission models: a mathematical framework. *Journal of the Royal Society Interface*, 9(68).
- Earnshaw, V. A., Eaton, L. A., Kalichman, S. C., Brousseau, N. M., Hill, E. C., & Fox, A. B. (2020). COVID-19 conspiracy beliefs, health behaviors, and policy support. *Translational Behavioral Medicine*.
- Farooq, A., Laato, S., & Islam, A. K. M. N. (2020). Impact of online information on self-isolation intention during the COVID-19 pandemic: cross-sectional study. *Journal of Medical Internet Research*, 22(5).
- Fast, S. M., González, M. C., Wilson, J. M., & Markuzon, N. (2015). Modelling the propagation of social response during a disease outbreak. *Journal of the Royal Society Interface*, 12(104).
- Finset, A., Bosworth, H., & Butow, P. (2020). Effective health communication - a key factor in fighting the COVID-19 pandemic. *Patient Education and Counseling*, 103(5).
- Friston, K. J. (2010). The free-energy principle: a unified brain theory? *Nature Reviews. Neuroscience*, 11(2), 127–138.
- Friston, K. (2013). Life as we know it. *Journal of the Royal Society Interface*, 10(86), 20130475.
- Friston, K. J. (2020). A free energy principle for a particular physics.
- Friston, K., Kilner, J., & Harrison, L. (2006). A free energy principle for the brain. *Journal of Physiology, Paris*, 100(1–3), 70–87.
- Friston, K. J., Daunizeau, J., Kilner, J., & Kiebel, S. J. (2010). Action and behavior: a free-energy formulation. *Biological Cybernetics*, 102(3).
- Friston, K. J., FitzGerald, T., & Rigoli, F. (2016). Active inference: a process theory. *Neural Computation*, 29, 1–49.
- Friston, K. J., Parr, T., & de Vries, B. (2017). The graphical brain: Belief propagation and active inference. *Nature Neuroscience*, 1(4), 381–414.
- Friston, K. J., Parr, T., Zeidman, P., Razi, A., Flandin, G., Daunizeau, J., Hulme, O. J., Billig, A. J., Litvak, V., Price, C. J., Moran, R. J., & Lambert, C. (2021). Second waves, social distancing, and the spread of COVID-19 across the USA. *Wellcome Open Research*, 5, 103.
- Funk, S., Salathe, M., & Jansen, V. A. A. (2010). Modelling the influence of human behaviour on the spread of infectious diseases: a review. *Journal of the Royal Society Interface*, 7(50), 1247–1256.
- Galang, C. M., Johnson, D., & Obhi, S. S. (2021). Exploring the Relationship Between Empathy, Self-Constraint Style, and Self-Reported Social Distancing Tendencies During the COVID-19 Pandemic. *Frontiers in Psychology*, 12, 588934.
- Golec de Zavala, A., & Federico, C. M. (2018). Collective narcissism and the growth of conspiracy thinking over the course of the 2016 United States presidential election: A longitudinal analysis. *European Journal of Social Psychology*, 48(7), 1011–1018.
- Hayward, A. C., Beale, S., Johnson, A. M., Fragaszy, E. B., & Group FW. (2020). Public activities preceding the onset of acute respiratory infection syndromes in adults in England - implications for the use of social distancing to control pandemic respiratory infections. *Wellcome Open Research*, 5(54).
- Hesp, C., Smith, R., Parr, T., Allen, M., Friston, K. J., & Ramstead, M. J. D. (2021). Deeply Felt Affect: The Emergence of Valence in Deep Active Inference. *Neural Computation*, 33(2), 398–446.
- Hu, Y., Wang, Y., Liang, H., & Chen, Y. (2017). Seasonal influenza vaccine acceptance among pregnant women in Zhejiang Province,

- China: evidence based on health belief model. *International Journal of Environmental Research and Public Health*, 14(12).
- Imhoff, R., & Lamberty, P. (2020). A Bioweapon or a Hoax? The Link Between Distinct Conspiracy Beliefs About the Coronavirus Disease (COVID-19) Outbreak and Pandemic Behavior. *Social Psychological and Personality Science*, 11(8), 1110–1118.
- Jang, W. M., Jang, D. H., & Lee, J. Y. (2020). Social distancing and transmission-reducing practices during the 2019 Coronavirus disease and 2015 Middle East respiratory syndrome coronavirus outbreaks in Korea. *Journal of Korean Medical Science*, 35(23).
- Janz, N. K., & Becker, M. H. (1984). The Health Belief Model: a decade later. *Health Education Quarterly*, 11(1).
- Jefferson, T., Del Mar, C., & Dooley, L. (2009). Physical interventions to interrupt or reduce the spread of respiratory viruses: systematic review. *British Medical Journal*, 339, b3675.
- Jiang, X., Elam, G., & Yuen, C. (2009). The perceived threat of SARS and its impact on precautionary actions and adverse consequences: a qualitative study among Chinese communities in the United Kingdom and the Netherlands. *International Journal of Behavioral Medicine*, 16(1).
- Kasperson, R. E., Renn, O., & Slovic, P. (1988). The social amplification of risk: A conceptual framework. *Risk Analysis*, 8, 177–187.
- Kassa, S. M., & Ouhinou, A. (2015). The impact of self-protective measures in the optimal interventions for controlling infectious diseases of human population. *Journal of Mathematical Biology*, 70(1–2).
- Khubchandani, J., & Macias, Y. (2021). COVID-19 vaccination hesitancy in Hispanics and African-Americans: A review and recommendations for practice. *Brain, Behavior, & Immunity – Health*, 15, 100277.
- Kissler, S. M., Tedijanto, C., & Goldstein, E. (2020). Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. *Science*, 5793.
- Kretzschmar, M., Rozhnova, G., Boven, M. (2021). Isolation and contact tracing can tip the scale to containment of COVID-19 in populations with social distancing. *Frontiers in Physics*.
- Laland, K. N. (2018). *Darwin's Unfinished Symphony: How Culture Made the Human Mind*. Princeton University Press.
- Lau, J. T., Yang, X., Tsui, H., & Kim, J. H. (2003). Monitoring community responses to the SARS epidemic in Hong Kong: from day 10 to day 62. *Journal of Epidemiology and Community Health*, 57(11).
- Lau, J. T., Tsui, H., Kim, J. H., & Griffiths, S. (2006). Perceptions about status and modes of H5N1 transmission and associations with immediate behavioral responses in the Hong Kong general population. *Preventive Medicine*, 43(5).
- Lau, J. T., Kim, J. H., Tsui, H., & Griffiths, S. (2007). Perceptions related to human avian influenza and their associations with anticipated psychological and behavioral responses at the onset of outbreak in the Hong Kong Chinese general population. *American Journal of Infection Control*, 35(1).
- Lau, J. T. F., Griffiths, S., Choi, K. C., & Tsui, H. Y. (2009). Widespread public misconception in the early phase of the H1N1 influenza epidemic. *The Journal of Infection*, 59, 122–127.
- Lee, M., & You, M. Psychological and behavioral responses in South Korea during the early stages of coronavirus disease 2019 (COVID-19). *International Journal of Environmental Research and Public Health*, 2020, 17–2977.
- Leung, G. M., Ho, L. M., & Chan, S. K. (2005). Longitudinal assessment of community psychobehavioral responses during and after the 2003 outbreak of severe acute respiratory syndrome in Hong Kong. *Clinical Infectious Diseases*, 40(12).
- Leung, G. M., Quah, S., & Ho, L. M. (2009). Community psycho-behavioural surveillance and related impact on outbreak control in Hong Kong and Singapore during the SARS epidemic. *Hong Kong Medical Journal*, 9, 30–34.
- Leung, C. C., Lam, T. H., & Cheng, K. K. (2020). Mass masking in the COVID-19 epidemic: people need guidance. *Lancet*, 395(10228).
- Liao, Q., Cowling, B. J., & Lam, W. W. T. (2014). Anxiety, worry and cognitive risk estimate in relation to protective behaviors during the 2009 influenza A/H1N1 pandemic in Hong Kong: ten cross-sectional surveys. *BMC Infectious Diseases*, 14(169).
- Lin, Y., Hu, Z., Alias, H., & Wong, L. P. (2020). Impact of mass and social media on psychobehavioural responses to COVID-19: A survey of medical university students in Fujian, China during the downward trend of COVID-19. *Journal of Medical Internet Research*, 10(2196/19982).
- Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welch, N. (2001). Risk as feelings. *Psychological Bulletin*, 127, 267–286. <https://doi.org/10.1037/0033-2909.127.2.267>
- Maddux, J. E., & Rogers, R. W. (1983). Protection motivation and self-efficacy: A revised theory of fear appeals and attitude change. *Journal of Experimental Social Psychology*, 19(5), 469–479.
- McCaskill, J. (2006). African-American women, self breast examination and Health Belief Model: implications for practice. *JOCEPS: The Journal of Chi Eta Phi Sorority*, 52(1), 33–37.
- Milne, S., Sheeran, P., & Orbell, S. (2000). Prediction and intervention in health-related behavior: A meta-analytic review of Protection Motivation Theory. *Journal of Applied Social Psychology*, 30(1), 106–143.
- Mizumoto, K., Chowell, G. (2020). Estimating Risk for Death from Coronavirus Disease, China, January–February 2020. *Emerging Infectious Diseases Journal*.
- Morgan, T. J. H., & Laland, K. N. (2012). The biological bases of conformity. *Frontiers in Neuroscience*, 6(87).
- Motta Zanin, G., Gentile, E., Parisi, A., & Spasiano, D. (2020). A Preliminary Evaluation of the Public Risk Perception Related to the COVID-19 Health Emergency in Italy. *International Journal of Environmental Research and Public Health*, 27, 17(9).
- Nields, J. A. (2020). Alone together in our fear: perspectives from the early days of lockdown due to COVID-19. *The Journal of Nervous and Mental Disease*, 208(6).
- Orbell, S., Perugini, M., & Rakow, T. (2004). Individual differences in sensitivity to health communications: consideration of future consequences. *Health Psychology*, 23(4).
- Person, B., Sy, F., & Holton, K. (2004). Fear and stigma: the epidemic within the SARS outbreak. *Emerging Infectious Diseases*, 10(2).
- Petrocchi, S., Bernardi, S., Malacrida, R., Traber, R., Gabutti, L., & Grignoli, N. (2021). Affective empathy predicts self-isolation behaviour acceptance during coronavirus risk exposure. *Scientific Reports*, 11(1), 10153.
- Piltch-Loeb, R., Abramson, D. M., & Merdjanoff, A. A. (2017). Risk salience of a novel virus: US population risk perception, knowledge, and receptivity to public health interventions regarding the Zika virus prior to local transmission. *PLoS One*, 12(12).
- Ramstead, M. J., Veissière, S. P., & Kirmayer, L. J. (2016). Cultural Affordances: Scaffolding Local Worlds Through Shared Intentionality and Regimes of Attention. *Frontiers in Psychology*, 7(1090).
- Ramstead, M. J. D., Badcock, P. B., & Friston, K. J. (2018). Answering Schrödinger's question: A free-energy formulation. *Physics of Life Reviews*, 24, 1–16.
- Ramstead, M. J., Kirchoff, M. D., Friston, K. J. (2019). A tale of two densities: active inference is enactive inference. *Adaptive Behavior*, 1059712319862774.

- Raude, J., Debin, M., Souty, C. (2020). Are people excessively pessimistic about the risk of coronavirus infection? PsyArxiv Preprint.
- Reicher, S., & Stott, C. (2020). On order and disorder during the COVID-19 pandemic. *The British Journal of Social Psychology*, 59(3).
- Rogers, R. W. (1975). A Protection Motivation Theory of Fear Appeals and Attitude Change. *The Journal of Psychology*, 91(1), 93–114.
- Rogers, R. W. (1983). *Cognitive and physiological processes in fear appeals and attitude change: a revised theory of protection motivation*. Guilford.
- Rosenstock, I. M. (1974). The health belief model and preventive health behavior. *Health Education*, 2, 354–386.
- Savadori, L., Rumiati, R., & Bonini, N. (1998). Expertise and regional differences in risk perception: the case of Italy. *Swiss Journal of Psychology*, 57, 101–113.
- Seale, H., Heywood, A. E., & Leask, J. (2020). COVID-19 is rapidly changing: Examining public perceptions and behaviors in response to this evolving pandemic. *PLoS One*, 15(6).
- Sesagiri Raamkumar, A., Tan, S. G., & Wee, H. L. (2020). Use of health belief model-based deep learning classifiers for COVID-19 social media content to examine public perceptions of physical distancing: model development and case study. *JMIR Public Health and Surveillance*, 6(3).
- Sharifirad, G., Yarmohammadi, P., Sharifabad, M. A., & Rahaei, Z. (2014). Determination of preventive behaviors for pandemic influenza A/H1N1 based on protection motivation theory among female high school students in Isfahan, Iran. *Journal of Education Health Promotion*, 3(7).
- Sharot, T. (2011). The optimism bias. *Current Biology: CB*, 21(23), R941–R945.
- Sharot, T., & Garrett, N. (2016). Forming Beliefs: Why Valence Matters. *Trends in Cognitive Sciences*, 20(1), 25–33.
- Slovic, P., Finucane, M. L., & Peters, E. (2004). Risk as analysis and risk as feelings: some thoughts about affect, reason, risk, and rationality. *Risk Analysis*, 24, 311–322.
- Steven, S., Yen Ting, L., & Chonggang, X. (2020). High Contagiousness and Rapid Spread of Severe Acute Respiratory Syndrome Coronavirus 2. *Emerging Infectious Diseases*, 26(7).
- Tang, C. S., & Wong, C. Y. (2003). An outbreak of the severe acute respiratory syndrome: predictors of health behaviors and effect of community prevention measures in Hong Kong, China. *American Journal of Public Health*, 93(11).
- Tang, C. S., & Wong, C. Y. (2005). Psychosocial factors influencing the practice of preventive behaviors against the severe acute respiratory syndrome among older Chinese in Hong Kong. *Journal of Aging and Health*, 17(4).
- Teasdale, E., Santer, M., & Geraghty, A. W. (2014). Public perceptions of non-pharmaceutical interventions for reducing transmission of respiratory infection: systematic review and synthesis of qualitative studies. *BMC Public Health*, 14(589).
- Turner, J. C., Hogg, M. A., & Oakes, P. J. (1987). *Rediscovering the social group: A self categorization theory*. Blackwell.
- Van Assche, J., Politi, E., Van Dessel, P., & Phalet, K. (2020). To punish or to assist? Divergent reactions to ingroup and outgroup members disobeying social distancing. *The British Journal of Social Psychology*, 59(3), 594–606.
- Veissière, S. P., Constant, A., & Ramstead, M. J. (2020). Thinking through other minds: A variational approach to cognition and culture. *The Behavioral and Brain Sciences*, 43, e90.
- Walker, K. K. (2021). Shining the Light on Altruism and Its Potential Impact for Behavior Change Strategies Related to the COVID-19 Pandemic and Beyond. *Health Communication*, 11, 1–3.
- Webster, R. K., Brooks, S. K., & Smith, L. E. (2020). How to improve adherence with quarantine: rapid review of the evidence. *Public Health*, 182, 163–169.
- Williams, L., Rasmussen, S., & Kleczkowski, A. (2015). Protection motivation theory and social distancing behaviour in response to a simulated infectious disease epidemic. *Psychology, Health & Medicine*, 20(7).
- Witte, K., & Allen, M. (2000). A meta-analysis of fear appeals: implications for effective public health campaigns. *Health Education & Behavior*, 27(5).
- Wolf, L. J., Haddock, G., Manstead, A. S. R., & Maio, G. R. (2020). The importance of (shared) human values for containing the COVID-19 pandemic. *The British Journal of Social Psychology*, 59(3).
- Xantus, G. Z., Burke, D., & Kanizsai, P. (2021). How to best handle vaccine decliners: scientific facts and psychological approach. *Postgraduate Medical Journal*.
- Xiao, H., Peng, M., & Yan, H. (2016). An instrument based on protection motivation theory to predict Chinese adolescents' intention to engage in protective behaviors against schistosomiasis. *Glob Health Res Policy*, 1(15).
- Xu, J., & Peng, Z. (2015). People at risk of influenza pandemics: The evolution of perception and behavior. *PLoS One*, 10(12), 0144868.
- Yamamoto, N., Jiang, B., & Wang, H. (2021). Quantifying compliance with COVID-19 mitigation policies in the US: A mathematical modeling study. *Infectious Disease Modelling*, 6, 503–513.
- Yousuf, H., Corbin, J., & Sweep, G. (2020). Association of a public health campaign about Coronavirus Disease 2019 promoted by news media and a social influencer with self-reported personal hygiene and physical distancing in the Netherlands. *JAMA Network Open*, 3(7).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.