ERC Consolidator Grant 2018
Research proposal [Part B1]

Improving collective decisions by eliminating overconfidence: mental, neural and social processes

rid-O

Cover Page:

- Name of the Principal Investigator (PI): **Bahador Bahrami**
- Name of the PI's host institution for the project: **Ludwig Maximillian University, Munich**
- Proposal duration in months: **60 months**

"What would I remove if I had a magic wand? Overconfidence" Daniel Kahneman's famous fairy-tale wish (*The Guardian* 18 Jul 15) conveyed a deeply seated pessimism in cognitive scientists that overconfidence is hardwired into human cognition. This bias is pervasive, costly and the root cause of many human failures. Previous failed attempts at reducing overconfidence targeted individual decisions. I propose to reduce this bias at social level of decision making by determining the underlying mental, neural and social processes involved in overconfidence and testing these models by causal interventions. I focus on 3 common forms of social decisions.

1. In honest communication of uncertainty (e.g. 2 doctors disagreeing over a diagnosis), overconfidence impairs joint decisions. Combining computational analysis of behaviour and brain response, I develop a real-time feedback loop that allows each disagreeing agent to weight her opinion by an individually-tailored signature of her own uncertainty, freeing joint decisions from overconfidence.

2. Focusing on advising and consulting, I use a novel laboratory model (i.e. Advising Game) to develop a theoretical and empirical understanding of overconfidence in the presence of conflict of interest to understand the mental and neural processes underlying strategic manipulation of others. Plus, by connecting the social use of overconfidence to self-esteem and self-worth, I translate this research to a Mental Health application looking at social dysfunctions in Depression.

3. Overconfidence impairs group processes (e.g. a panel selecting among grants) by promoting herding (blindly following others) and polarization to extreme viewpoints. Inspired by a recent discovery from my lab, I develop a novel causal intervention that puts together seeking consensus within each group and aggregating consensus opinions across groups to remove herding and polarization.

Collective decisions can be made better and rid-O could help this wish come true.

1. **State of the Art and Objectives: What is overconfidence? Why does it matter?**

Human decision making is flawed by numerous biases. Among them, overconfidence is perhaps the most pernicious one. Overconfident Theresa May called a disastrous snap election in UK in 2017. Only a year before, overconfident David Cameron had called the even more disastrous 2016 EU referendum. Politicians are not the only victims of overconfidence: we have a very good intuitive estimate of the chance that a colleague (who is in our age and health condition) may die of cancer. But we grossly underestimate the same probability when thinking about ourselves. Another example: the overwhelming majority of restaurants that open in London go bankrupt within 1 year; still, many new ones open every day. Overconfidence is pervasive and costly.

Without overconfidence, however, almost all human ambitions (including the high-risk high-reward ERC-COG-2018 applications) should be rejected, grinding human civilization to a halt of thoughtfully measured judiciousness. Overconfidence may sometimes be inevitable or even beneficial. Some suggest that overconfidence might even be hardwired into human cognition and not easily eradicated. Removing overconfidence entirely from human cognition is not just hard but probably harmful too.

Overconfidence covers a range of biases extending from individual to social. Optimism bias makes us learn from and expect good (but not bad) news[1]. Confirmation bias refers to interpreting all previous experiences in our favour [2]. More socially, we overrate our competence compared to others (Dunning-Kruger effect) and [3] Equality Bias refers to the unwillingness to accept that some (e.g. experts) may know better than others [4]. I suggest that previous attempts to reduce overconfidence failed because they focused on individual rather than the social decision-making where it most relevantly occurs. The key question is to understand and (from there) reduce the negative impact of overconfidence on social interactive decision making.

What is interactive decision making? How is it related to overconfidence?

Interactive decision making is ubiquitous to everyday human life and comes in many forms (Fig 1). One form happens when, for example, two doctors disagree about a diagnosis or two football referees decide if there was a goal or not. These decisions are best made when disagreeing opinions are weighted by their respective reliability [5-7].

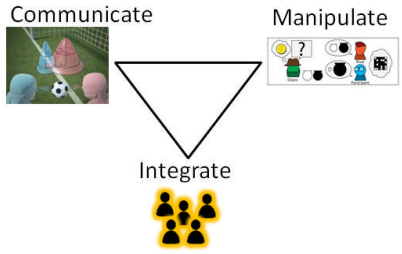

Figure 1. Three aspects of overconfidence shape the key questions in this proposal.

The problem is how the referees should communicate their uncertainty to one another. We use confidence to do this [8,9]. But confidence also varies with irrelevant psychological, social, cultural and demographic factors [10-13]. This makes inferring others' reliability a complex and demanding cognitive process.

To solve this problem, human mind resorts to a simplifying heuristic called Equality Bias [4,6,14]: when making joint decisions, we ignore inter-individual differences in reliability and behave *as if* everyone's is equally reliable. This is useful when decision partners have similar levels of expertise. It is unhelpful when partners have very different expertise: overconfidence of the lesser able partner damages joint decisions. This bias is universal among cultures[4,15] and robust to performance feedback and financial incentives [4]. In addition to simplifying cognition, equality bias also serves social functions such as maintaining inclusiveness [14] and sharing of responsibility for difficult decisions [16] explaining why technical solutions[4] that neglect these human and social factors are ineffective.

Thus, even when communicating effectively and honestly is in everyone's best interest, overconfidence still impairs joint decisions through equality bias[4,14,15]. To remove overconfidence in such situations, we should find a way to estimate our own uncertainty accurately before communicating it to others.

Another form of interactive decisions happens when a customer weighs the advice of a salesman or a voter considers supporting a politician. Unlike the first case, interests may conflict here [17,18] and overconfidence becomes far more complicated problem [19]. Baffled by the abominable public behaviour of overconfident politicians, we witness (with indignant surprise and disbelief) their relentless public success despite violating all social norms. Our confusion is perhaps due to the fact that cognitive sciences have told us so little about the mental and neural processes involved in *attempting* to influence others. In fact game theoretic analyses indicate that overconfidence does *promote* individual's interests at the expense of others; but we know very little about how the human mind and brain uses overconfidence to manipulate others.

Literature of social influence has– almost exclusively – focused on the mind and brain of the ***target*** of social influence (e.g. voter, consumer)[20-23]. The cognition at the ***source*** (e.g. politician, salesman) of influence has received far less attention. Such *source-target* mode of social interaction is not restricted to politician-voter but extends to parent-child, doctor-patient, teacher-student and numerous other social roles. For the consumer speaking to the salesman, deciding better involves combining her own uncertainty (see first

problem) with the advice from salesman after having separated the facts from the spin. To help the client achieve this and fend off strategic overconfidence, we need to understand the mind and brain of the adviser.
The third type of interactive decision involves aggregating of opinions across members of a group (e.g. a panel selecting grant proposals). We believe that crowds are wiser than individuals [24-30]. But, outside of social psychology laboratory, poor collective judgment is very common[5,31]. Interactions in a group perpetuate the communicative and manipulative (Fig 1) forms of overconfidence [32] that we saw above leading to herd mentality, groupthink, echo chambers and polarization[33], which in turn could make us support absurd or obviously wrong/immoral propositions[6,34]. Misled by overconfident individuals [35] human crowds are prone following each other blindly into to catastrophes such as financial bubbles [36] and polarizing to extreme positions that lead to violence contagion [37]. Many collective failures are rooted in overconfidence. The problem here is how to organize group interactions to benefit from "wisdom of crowds" without suffering from the group-level consequences of overconfidence (herding and polarization).

OBJECTIVES
Work package 1: develop an individually-tailored computational and neural signature of uncertainty, create a real-time feedback loop to weight disagreeing opinions by their uncertainty signature and thereby reduce equality bias and improve joint decisions.
Work package 2: use Advising Game as a laboratory model to develop a theoretical and empirical understanding of overconfidence in the presence of conflict of interest to understand the mental and neural processes underlying strategic manipulation of others.
Work Package 3: develop a systematic method to remove overconfidence from group processes by elucidating the role of seeking consensus within- and aggregation of opinions between groups.

2. Innovation and Impact

Novelty	-- In psychology, neuroscience and economics labs, decision confidence is studied in isolated individuals making private decisions. **Rid-O** studies overconfidence in *social domain* combining psychological, computational, game theoretic and neurobiological approaches. -- **Rid-O'**s looks for *solutions* to the problem of overconfidence through *causal interventions* that are informed by (and serve as critical tests of) the recent developments in social cognitive neuroscience. -- **Rid-O** diverges from previous works in social influence by investigating the cognition at the "source" of influence (eg salesman, politician) rather than focusing on its target (costumer, voter). -- **Rid-O** connects the decision confidence to concepts of self-esteem and self-efficacy and translates this research to a *Mental Health application*. -- **Rid-O** is the first to systematically examine benefits of aggregation of knowledge across groups producing the entirely new concept of *wisdom of structured crowds*. -- **Rid-O** experiments will be performed *cross-culturally* in a network of collaborations in Europe, Iran, Israel, China and Argentina.
Impact	-- For the mind and brain sciences, WP1-2 will identify the cognitive and neural basis confidence as a tool for social influence. New causal interventions (e.g. neurofeedback, group organization) will be developed to test these cognitive models. -- For mental health, WP2 will transform our understanding of social dysfunctions in depression by connecting self-esteem, confidence and social influence. For social and organizational psychology, management and economics, WP3 will develop new methods for efficient knowledge aggregation. -- For public, **rid-O** will help communicate uncertainty more effectively by avoiding equality bias; be better persuaders by using effective game theoretic strategies; protect ourselves better against unwelcome influence (e.g., from exuberant politicians); understand how our influence on others shapes our self-esteem and make more efficient and more accurate group decisions.

3. Methodology:

Work package 1 - overconfidence in communication of uncertainty

As we saw, Equality Bias arises from inter-individual differences in how we express our confidence. To remove this bias, I will employ a cognitive and computational framework (Fig 2) I developed for understanding inter-individual differences in confidence [8,38]. The participants observe a sequence of visual or numerical stimuli and decide about their perceived mean of the sequence and report their confidence. The variance of the sequences changes across trials. Using a simple learning model, we estimate (on a trial by trial basis) the participant's subjective probability of being correct (P_c, Fig 2 shaded area in bottom panel) separately from the perceived variance (σ^2, Fig 2) of the sequence. We then disentangle the impact of P_c and σ^2 on confidence. A large portion of between-individual variability in participants' confidence is explained by how a participant combines P_c and σ^2 to report her confidence. This combination is like a computational fingerprint: unique to each participant and stable across time and cognitive domains. Knowing how your confidence relates to P_c in averaging numbers today would help interpret your confidence when averaging

orientations in two weeks' time [38]. **If this fingerprint informs joint decision making, inter-individual differences in reliability become readily apparent and equality bias can be eradicated.** These hypotheses operationalize this idea.

Hypothesis H1 Equality bias will be reduced and collective decision making will be more successful if people making joint decisions use the computational fingerprint procedure to adjust their confidence to their level of expertise.

Hypothesis H2 The computational fingerprint of confidence has identifiable neural correlates that can be dissociated and read out from electroencephalography (EEG) signal recorded from human scalp.

Hypothesis H3 Combining H1 & H2, if joint decision making is provided with a real-time feedback of this personalized computational and neural fingerprint, collective decision will be substantially improved.

To test H1, I will combine my Interacting Mind [39] and Computational Fingerprint [38] paradigms to construct an intervention to eradicate overconfidence from joint decision making. Pairs of participants will make first make individual perceptual decisions (cf Fig 2D and ref [38]). Then they communicate their confidence to reach a joint decision. From this control condition, we estimate participants' individual ability, their level of equality bias and the computational fingerprint of their confidence which tells us how the person's confidence relates to her perceived probability of being correct [38]. In the key intervention (Fig 3E), after making an individual decisions, the subject gets feedback about her own model-based probability of correct choice. She could then *choose to adjust* her confidence before communicating it to their partner. Prediction: joint decisions will be more accurate and equality bias will be reduced in the intervention vs control condition.

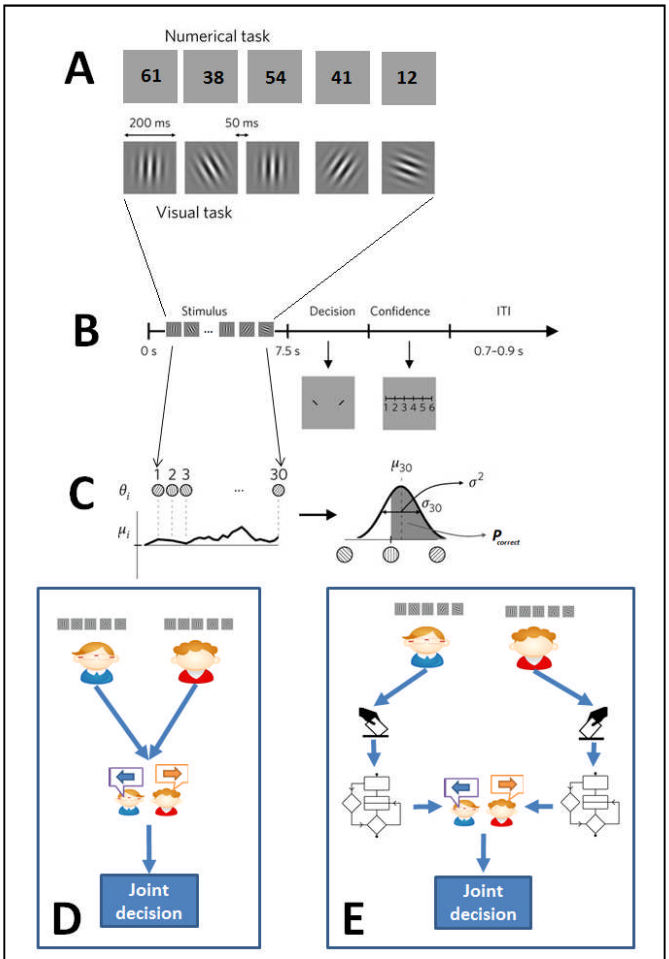

Figure 2. (A,B) Experimental Paradigm. Stimulus: a rapid (4Hz, duration:7.5s) sequence of 30 visual stimuli (numbers or oriented gratings). Task: compare the mean of the observed sequence to a boundary and report Confidence on a 1-6 scale. (C) A computational model estimates the P_c (shaded area in the Gaussian) and variance (σ^2) for each trial. (D) Combining Interacting Minds paradigm (Bahrami et al *Science* 2010) with Computational Fingerprinting (Navajas et al *Nat Hum Behav* 2017). (E) H1: Interactive decision making via model-based (symbolized by flowchart), subject-specific advice.

I will test H2 by combining my Computational Fingerprint [38] paradigm with electroence-phalography (EEG) to disentangle the neural correlates of subjective probability and perceived variance. Importantly, the visual stimuli (high contrast gratings, 7.5s at 4Hz, Fig 2A) are optimized for inducing steady state visual evoked potential (SSVEP) [40] to create a strong EEG signal. I will use multivariate time-frequency EEG analysis [41-43] and trial by trial model-based analysis of behaviour. Predictions: (1) neural correlates of perceived variance (σ^2) will be stronger EEG power over the occipital and parietal cortex in trials with low (vs high) stimulus variance independent of accuracy. Correlates of subjective probability (P_c) will be identified by an interaction: EEG power in trials with low stimulus variance will be higher for correct vs error trials. Our pilot data (N=7; see Part B2) show very promising results for timing and scalp distribution of the neural fingerprint of confidence.

I will test H3 by combining the procedures developed in H1 and H2 and conduct a composite experiment involving dyadic interactions and real-time EEG feedback. Pairs of participants will make individual and dyadic decisions. EEG will be recorded from both participants. H1 will be used to infer the psychometric probability of correct choice (pP_c) [38]. Following H2, EEG will be analysed in real-time to produce a trial-by-trial neurometric probability of correct choice. Each partner will then be advised about her own nP_c and/or pP_c before discussing joint decision. Prediction: combining the nP_c and/or pP_c will reduce equality bias dramatically, leading to the highest accuracy in collective decisions and removing overconfidence.

Work Package 2: Overconfidence in strategic manipulation

I recently developed [18] a novel laboratory model for giving advice (Fig 3; web demo http://bit.ly/2lhh0Vk) that, for the first time investigates the cognition underlying attempting to influence others. Groups of 3 participants play an Advising Game in the lab (Fig 3). The Client wants to buy a lottery ticket but doesn't know which lottery (black or white) is better. Two other participants (i.e. the Advisors) privately receive facts about each lottery's probability of winning and advise the client publically by expressing their choice of lottery and confidence. Preliminary findings showed that advisors' overconfidence depends on her social status (is the client ignoring the advisor?) and self-image (does advisor believe she is better than her rival?). Importantly, this paradigm permits disentangling the contribution of objective facts and strategic motives to confidence. Armed with this laboratory model, using a multidisciplinary approach that combines theoretical, behavioural and computational and neuroimaging methods, I will investigate overconfidence in advising with theoretical and empirical questions.

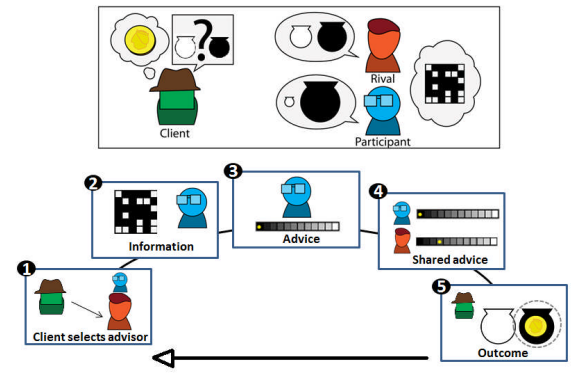

Figure 3. Advising Game

In collaboration with Stephan Hartmann and Ophelia Deroy (LMU) **I will formalize exerting social influence as a game theoretic problem to determine what behaviour maximises adviser's influence in the advising game.** For advisers, competition over influencing the client is a zero-sum game: one adviser's success is the other's failure. Employing the methods developed to solve zero-sum game (e.g. backward induction [44]), I will create a theoretical benchmark to interpret the advising behaviour with.

Empirically, I will investigate the role of (1) rival disposition, (2) client knowledge, (3) ambiguity of facts, (4) rewards and (5) punishments in strategic overconfidence; I will use (6) reinforcement learning to interpret the behaviour. For brevity, here I describe three of these hypotheses (2 behavioural and 1 fMRI) that form one complete idea together (for a complete description of the other hypotheses see Part B2).

Hypothesis H1 Advisers' overconfidence closely tracks their rival's disposition.

Hypothesis H2 Advisers' overconfidence depends on their belief about the client's knowledge.

Hypothesis H3 Distinct brain areas track self- and rival's performance. A network of brain areas - involved in Theory of Mind [45] - tracks the belief about Client's knowledge. All of these brain areas interact to shape overconfidence.

To test H1, in the Advising Game, I will pair the participants with (1) overconfident or (2) underconfident rivals. I predict that working against an overconfident rival will exaggerate the participant's overconfidence.

To test H2, I will compare adviser's overconfidence when client is (1) naïve (has no information about lottery chances – control condition); (2) informed: the client has some information but is asking for a second opinion. Importantly, I will keep the client's behaviour identical across the two conditions and only the advisor's belief about client's knowledge will be manipulated. Prediction: advisors will more overconfident towards the client who is declared as naïve.

To test H3, in a repeated measure 2x2 blocked design, I will test the neuronal substrates of how adviser overconfidence is shaped by beliefs about rival disposition (H1) and client knowledge (H2). Behavioural methodology will combine H1 and H2. Prediction: human right temporo-parietal junction (rTPJ) and bilateral mid-cingulate cortex [46] will track the main effect of belief about client knowledge. The main effect of rival's behaviour will be observed in dorsomedial prefrontal cortex (dmPFC) [47].

Why fRMI? fMRI is the method of choice to test H3 prediction of activity in deep brain structures (eg Ventral Striatum) that require high spatial resolution.

Advising in mental health. In collaboration with Leonhard Schilbach (Max Plank Psychiatry - Munich) **I will use the Advising Game to translate my basic research to a new mechanistic understanding of depression**. Preliminary findings[18] show that influencing others affects our sense of well-being and social anxiety. The hypothesis is that influencing strategies are impaired in depression. I test whether the model-based estimates of social comparison and social influence drawn from Advising Game [18] correlate with clinical severity of depressive symptoms. Our promising pilot results (see Part B2) point to a new neuro-computational hypothesis for deterioration of self-esteem in depression.

Work Package 3. Overconfidence in integration across crowds

Recently, we ran a field experiment [48] to find ideas for how to improve collective accuracy. We asked a live crowd ($N = 5{,}180$) gathered in a stadium to respond to general-knowledge questions (e.g., "What is the height

of the Eiffel Tower?"). Participants first answered individually, then deliberated to reach consensus estimate in groups of five.Remarkably, combining as few as 4 consensus estimates outperformed the wisdom of thousands of individuals. Analysis of the observational data suggested that locally structured deliberation decreased the diversity of opinions *within* each small group, similar to "herding" and increased *between*-group distance, similar to what is seen in *polarization*. These descriptive observations are remarkable but they do not help us understand the psychological and cognitive *mechanisms* involved in how aggregated opinions from 4 groups of 5 could beat 5000 individuals. They do however, raise a counterintuitive causal hypothesis: **if seeking consensus within groups is combined with aggregating them across groups, the two consequences of overconfidence (i.e. herding and polarization) cancel each other out and collective decisions improve.** Here, I describe 2 hypotheses to test this mechanistic explanation. Part B2 includes another, related experiment.

Hypothesis H1 Commitment to reaching a consensus in a small group increases herding *within* the groups and polarization *between* groups.

Hypothesis H2 Commitment to reaching a consensus within small groups increases wisdom of the crowds *across* groups.

I will test H1 and H2 by replicating the experimental conditions of our original field study [48] in the lab. I will compare two conditions: (1) groups members are required to reach a consensus through discussion; (2) groups discuss their opinions but are not required to reach a consensus. These predictions are extremely counterintuitive: numerous studies have demonstrated the detrimental effect of social influence (i.e. seeking consensus) on wisdom of crowds [34,49]. I contend that those previous studies replaced face-to-face interaction with connecting people via virtual digital networks, which are methodologically convenient but psychologically implausible and perhaps inadequate replacements for social interaction.

Risk Management: WP1 consists of high-risk, high-return experiments both in terms of technical complexity as well as expected results. Anticipating these challenges, I have taken a number of measures to meet them. The experimental paradigm is designed (and confirmed by pilot data) to produce robust EEG signals. I start rid-O with this project to make sure adequate time and resources can be dedicated to it. I will collaborate with a team of international experts led by Prof. Herman Muller at LMU. WP2 combines safer experiments in behaviour and imaging with a risky translational project in Mental Health. My past experience as a medical doctor (MD thesis in mood disorders) will help me deliver this project. The experiments in WP3 are safe with simple experimental paradigm and predictions.

	2019				2020				2021				2022				2023			
	I	II	III	IV	I	II	III	IV	I	II	III	IV	I	II	III	IV	I	II	III	IV
WP1 – communication																				
H1 – Behaviour	p1	p1	p1	p1																
H2 – EEG			p1	p1	p1	p1														
H3 - combined Beh & EEG							p1	p1	p1	p1	p1	p1								
WP2 – manipulation																				
Game Theory	pi	pi	pi	pi	pi	pi	pi	pi	pi	pi	pi	pi	pi	pi						
H1,2 – Behaviour					p2	p2	p2	p2	p2	p2	p2	p2								
H3,4 – fMRI									p2	p2	p2	p2	p2	p2	p2	p2				
Mental Health	ra	ra	ra	ra	ra	ra	ra	ra	ra	ra	ra	ra	ra	ra	ra	ra	ra	ra	ra	ra
WP3 – aggregation																				
H1,2 – Consensus									p3	p3	p3	p3	p3	p3	p3	p3				
H3 – Group size (see B2)											p3	p3	p3	p3	p3	p3	p3	p3	p3	p3

Table 1. Timeline of proposed activities. **pi** indicates work directly assigned to principle investigator. **p1-p3** refers to postdoc 1-3. **ra** refers to research assistant.

Sample Size Estimation: A table of sample size estimates and effect sizes is provided in Part B2.

4. Resources.

Why LMU? I have recently moved my research from UK to Germany and taken an academic position at Ludwig Maximillian University in Munich. LMU hosts a concentration of experts in experimental psychology (Hermann Muller, for the EEG experiments), neuroscience (Benedict Grothe), philosophy (Ophelia Deroy, Stephan Hartmann), and economics (Klaus Schmidt) who are my collaborators. Proximity to Max Plank Institute of Psychiatry hospital (collaborating with Leonhard Schilbach) provides access to MRI facilities and mental health clinic. My host department guarantees support of my research program and career development as well as space and resources for the tenure of the award. I will dedicate 70% of my time to rid-O project.

Personnel (70%PI + 3 Postdocs, 1RA)	1,322,130
Travel	25,000
Other Direct Costs (MRI, subj payment etc)	221,000
Indirect Costs (overheads) 25% of Direct Costs	385,782.50
Total Requested Budget (€)	**1,928,912.50**

Brief justification of costs. Salary is requested for 70% of PI time, three full-time graduate postodcs and one part-time (50%) research assistant. Each postdoc is assigned to one work package (see Table 1). One part-time research assistant is assigned to the mental health project. Travel funds (at economy rate) to international conferences requested for the PI (€2K per year) and postodcs (€1.67K per year per each postdoc). Other direct costs include (1) costs of open access publication (€5x10K) and (2) subject compensation (5x6K € for testing a total of N=3000 hour- participant of at a rate ~10.00€ per hour) and (3) fMRI imaging costs (50K €) performed in Max Planck Institute of Psychiatry Hospital (for a total of 125 hours at the hourly rates which start from 400€).

Section b: Curriculum vitae

PERSONAL INFORMATION

Bahrami, Bahador

ORCID: 0000-0003-0802-5328

Date of birth: 16 Apr 1977

Nationality: United Kingdom, Iran

URL for web site: http://bit.ly/2bO7nuD

- **EDUCATION**

2008 PhD (topic: Attention and awareness in human behaviour and brain)
Institute of Cognitive Neuroscience, University College London, UK
PhD Supervisor: Vincent Walsh

2003 Medical Doctorate (MD)
Department of Medicine, Tehran University of Medical Sciences, Iran
Thesis: Resting state brain activity in mood disorders

- **CURRENT POSITION(S)**

2007 – Present Senior Research Fellow in *Social Cognition* (Fulltime)
Faculty of Psychology and Educational Sciences, Ludwig Maximilian University, Munich, Germany

2013 – Present Group Leader at *Crowd Cognition Lab* (from Jan 2018 as Honorary)
Institute of Cognitive Neuroscience, University College London, UK

- **PREVIOUS POSITIONS**

2010 – 2012 British Academy Postdoctoral Researcher
Institute of Cognitive Neuroscience, University College London, UK

2009 – 2010 Postdoctoral Research Fellow
Interacting Mind Centre, Aarhus University, Aarhus, Denmark

- **FELLOWSHIPS**

2013 – 2017 European Research Council Starting Grant, Institute of Cognitive Neuroscience, University College London UK

2010 – 2012 British Academy Postdoctoral Fellow, Institute of Cognitive Neuroscience, University College London, UK

2005 – 2006 Sully Scholarship, Institute of Cognitive Neuroscience, University College London, UK

2004 – 2007 Overseas Research Scholarship (ORS), Institute of Cognitive Neuroscience, University College London, UK

2004 – 2007 Graduate School Research Scholarship, Institute of Cognitive Neuroscience, University College London, UK

- **SUPERVISION OF GRADUATE STUDENTS AND POSTDOCTORAL FELLOWS**

2013 – 2017 4 Postdocs/ 2 PhD/ 10 MSc Students
Institute of Cognitive Neuroscience, University College London, UK

2010 – 2017 2 PhD/ 12 MSc Students
Department of Artificial Intelligence and Electrical Engineering, Tehran University Iran

2009 – 2010 4MSc Students
Interacting Mind Centre, Aarhus University, Aarhus, Denmark

- **TEACHING ACTIVITIES**

2010 – 2017 Lecturer – Experimental Methods in Cognitive Sciences, Tehran University Iran

2013 – 2017 Lecturer – MSc program in Cognitive Neuroscience, Institute of Cognitive Neuroscience, University College London UK

2008 – 2017 Technical Instructor – Human Structural Brain Imaging, School of Cognitive Sciences, Iranian Centre for Theoretical Physics and Mathematics Iran

- **ORGANISATION OF SCIENTIFIC MEETINGS**

2016 Course Organizer – IBRO Regional Summer School in Human Brain Imaging, >50 attendees, Cape Town, South Africa

2015 Course Organizer – IBRO Regional Summer School in Human Brain Imaging, >100

attendees, Tehran, Iran

2012-2-13 Course Organizer (in collaboration with Prof Patrick Haggard) - IMPRS NeuroCom Summer School, London UK

- **INSTITUTIONAL RESPONSIBILITIES**

2013– Group Leader, Institute of Cognitive Neuroscience, University College London, UK

2015– Graduate Student Advisor, Department of Economics, Sharif Institute of Technology, Tehran, Iran

2013– Graduate Student Advisor, Department of Artificial Intelligence and Electrical Engineering Tehran University, Iran

2012 – 2013 Organizer of FIL Brain Meetings, Welcome Centre for Neuroimaging, University College London, UK

- **COMMISSIONS OF TRUST**

2014 – Scientific Advisory Board, Cognitive Sciences and Technologies Council, Tehran Iran

2015-2016 Student Selection Panel, International Brain Research Organization (IBRO)

2004 – Ad-hoc peer review for Journal of Neuroscience, Proceedings of National Academy of Sciences, Nature Human Behaviour, Current Biology, Journal Of Experimental Psychology: General; Journal Of Experimental Psychology: Human Perception and Performance; Psychonomic Bulletin and Review, Cerebral Cortex, Journal of Cognitive Neuroscience, Perception

- **MAJOR COLLABORATIONS**
  - International Brain Research Organization (IBRO) – technology transfer and training programs in Iran (2015) and South Africa (2016)
  - Gatsby Computational Unit, UCL London (Peter Latham)
  - Munich Centre for Neuroscience, Ludwig Maximilian University, Munich, (Ophelia Deroy)
  - Universidad Torcuato Di Tella, Buenos Aires and TED Argentina (Joaquin Navajas)
  - Nippon Telegraph and Telephone, Atsugi, Japan (Shiro Kumano)
  - School of psychological and cognitive sciences, Peking University (Jian Li)
  - Faculty of Electrical Engineering, Tehran University, Iran (Majid Nili Ahmadabadi)
  - Departement d'Études Cognitives, École Normale Supérieure, Paris, (Stefano Palminteri)
  - Max Planck Centre for Mental Health & Aging, UCL, London (Raymond Dolan)
  - Interacting Mind Center, Arhus University, Aarhus Denmark (Andreas Roepstorff)
  - Araya Brain Imaging Inc., Tokyo Japan (Ryota Kanai)
  - Centre for Adaptive Rationality, Berlin, Germany (Ralph Hertwig)
- **Outreach and public engagement**

2017 Thought Leadership Forum event held by Credit Suisse, Vienna Austria

2017 World Economic Forum, Davos Switzerland http://bit.ly/2v7x0OM

2016 South African Society for Neuroscience meeting Cape Town South Africa

2015 Financial Education and Investor Behaviour at Rio de Janeiro, held by the Brazilian Securities Exchange Commission (CVM)

Appendix: All ongoing and submitted grants and funding of the PI (Funding ID)

Mandatory information (does not count towards page limits)

Previous Grants

Project Title	*Funding source*	*Amount (Euros)*	*Period*	*Role of the PI*	*Relation to current ERC proposal*
Neurobiological basis of collect-ive decision making in the human brain	ERC-StG	1,486,197	5yrs – (Jan 13- Dec 17)	Principle Investigator	Current proposal addresses different questions and goes beyond the previous ERC in several significant respect
Are two heads better than one? The cognitive neuroscience of collective decision making	British Academy	282,000	2 yrs (Sept 10- Dec 12)	Principle Investigator	Formed the basis of my current research program

On-going Grants

Project Title	*Funding source*	*Amount (Euros)*	*Period*	*Role of the PI*	*Relation to current ERC proposal*[1]
An adaptive role for collective deicions: shared responsibility in the human brain	Sir Henry Wellcome Postdoctoral Fellowship, UK	283,340	4 yrs (Jan 17-Dec 20)	Mentor	Complimentary but not overlapping. The project looks at role of responsibility in joint decisions
Deep brain communication science project II	Nippon Telegraph and Telephone, Japan	30,260	3 yrs (Aug 16-Dec 19)	Collaborating PI	Complementary but not overlapping. The project focuses on understanding the human facial emotional reaction in social interaction

Grant applications

Project Title	*Funding source*	*Amount (Euros)*	*Period*	*Role of the PI*	*Relation to current ERC proposal*[2]
Our influence on others: from cognitive neuroscienc e to mental health	Senior NonClinical Fellowship (SNCF) at Medical Research Council (MRC) UK	1,504,670.00	5yrs (July 18 – Jun 23)	Principle Investigator	This project overlaps with work page #2 of the ERC proposal. The research questions and objectives are similar. The project has a strong translational component for mental health.

[1] Describe clearly any scientific overlap between your ERC application and the current research grant or on-going grant application.

					In the case of receiving the ERC-CoG and SNCF, I will resign from SNCF fellowship to dedicate my time to ERC-CoG.
Wisdom of Structured Deliberating Crowds	Humboldt Research Fellowship for Experienced Researchers	Unspecified by the grant call. Salary commensurate with qualifications will be provided for the period of fellowship (~13months)	18 mths (Sep 18 – Dec 19	Principle Investigator	This project overlaps with work page #3 of the ERC proposal. The research questions and objectives are similar. The project has a strong translational component for management and organizational sciences. In the case of receiving both the ERC-CoG and this fellowship, I will resign from this fellowship to dedicate my time to ERC-CoG.

Section c: Early achievements track-record (max. 2 pages)
(✤ Main Senior Author) (✍ : Main Lead Author) (→ Without PhD supervisor)
Total number of publications (by end of Jan 2018): 61
google scholar metrics: Total citations: 3452; h-index: 31; i10-index: 47

1. → Navajas, J., Niella, T., Garbulsky, G., Bahrami, B., & Sigman, M. (2018) Aggregated knowledge from a small number of debates outperforms the wisdom of large crowds. *Nature Human Behavior*
   Describes our first findings about how to enhance collective decision making through hierarchical crowd structure combined with deliberation and consensus.
2. → ✤ Hertz, U., Palminteri, S., Brunetti, S., Olesen, C., Frith, C.D., & Bahrami, B. (2017) Neural computations underpinning the strategic management of influence in advice giving. *Nature Communications* 1(8):2191.
   Whereas previous works on social influence studied the TARGET of influence (e.g. customer, voter), this paper combines game theory, computational behavioural analysis and fMRI to understand the mind and brain of the SOURCE of the influence (i.e. salesman, politician, consultant).
3. → ✤ Navajas, J., Hindocha, C., Foda, H., Keramati, M., Latham, P. E., & Bahrami, B. (2017) The idiosyncratic nature of confidence. *Nature Human Behaviour*.
   People vary widely in how they communicate uncertainty. This paper offers a computational framework for explaining these inter-individual differences by producing a computational fingerprint for each participant that is stable across time and cognitive domains.
4. → Bang, D., Aitchison , L., Moran, R., Castanon S.H., Rafiee, B., Mahmoodi A., Lau J.Y.F., Latham, P.E., Bahrami, B & Summerfield, C. (2017) Confidence matching in group decision-making. *Nature Human Behaviour*1:0117.
5. → ✤ Pescetelli, N., Rees, G., & Bahrami, B. (2016) The perceptual and social components of metacognition. *Journal of Experimental Psychology: General.* 145(8):949-65.
6. → Aitchison L, Bang D, Bahrami B & Latham PE (2015) Doubly Bayesian Analysis of Confidence in Perceptual Decision-Making. *PLoS Comput Biol* 11(10): e1004519.
7. → ✤ Mahmoodi, A., Bang, B., Olsen, K., Zhao, Y.A., Shi, Z., Broberg, K., Safavi, S., Han, S., Ahmadabadi, M.N., Chris, D.F., Roepstorff, A., Rees, G., & Bahrami, B. (2015) Equality bias impairs collective decision-making across cultures. *PNAS* 12, 3835–3840.
   To make the best joint decisions, disagreeing opinion must be scaled according to its reliability. This paper showed that in Denmark, Iran, and China, people do not follow this sensible strategy: they treat all opinions as equal regardless of their reliability, even despite explicit feedback or monetary incentives.
8. → Fusaroli, R., Bahrami, B, Olsen, K., Roepstorff, A., Rees, G., Frith, C., & Tylén K. (2012) Coming to Terms: Quantifying the Benefits of Linguistic Coordination. *Psychological Science.* 23(8):931-9.
9. → ✍ Bahrami, B., Olsen, K., Latham, P., Roepstorff, A., Rees, G., & Frith, C. (2010). Optimally interacting minds. *Science*. 329: 1081-1085.
   Are two heads really better than one? This paper cast this old problem in a new conceptual and computational framework that combined sensory psychophysics, computational neuroscience and social psychology.
10. ✍ Bahrami, B**.** Lavie, N. & Rees, G. (2007) Attentional load modulates responses of human primary visual cortex to invisible stimuli. *Current Biology* 17, 509-513.

Grants, Awards & Honours (see page 16 for previous grants)

2005-2006	Sully Scholarship, UK
2005	Oliver Braddick Prize in Psychology
2004-2007	Overseas Research Scholarship (ORS), UK
2004-2007	Graduate School Research Scholarship (GSRS), UK

Achievements

In 1996, when I was an undergraduate medical student in Iran, together with a group of my friends we founded an independent student research group that slowly gained international recognition culminating in a news story published in ***Nature* in 2005** http://bit.ly/ru2Xkj.
In my career, I have been committed to contributing to knowledge transfer to developing countries. I have regularly taught empirical methods in neuroscience at regional schools in Iran and once in South Africa. For the past 6 years I have held a bi-weekly web-based journal club connecting students from several cities around Iran. This network of students has now produced 6 papers in international peer reviewed journals.
I have put together an international network of scientists, laboratories and companies that all share an interest in cross-cultural examination of collective decision making. This network has allowed my research to

integrate data from Japan, China, Iran, Israel, Denmark, France, Germany, Italy, UK and most recently, Argentina.

Training and supervision

In my career I have supervised more than 25 MSc and MA students. I have also contributed to the supervisions of 15 PhD students (examples include: Nahid Zokaei currently a British Academy postdoctoral fellow at university of Oxford; Dan Bang postdoctoral researcher at UCL; Nima Khalighinejad postdoctoral researcher at University of Oxford; Ali Mahmoodi PhD candidate at Freiburg University). Two postdocs (Uri Hertz and Joaquin Navajas) who trained in my lab have now gone back to their home countries in Israel and Argentina in prestigious academic positions. I mentored Marwa El Zein to obtain a Sir Henry Wellcome postdoctoral fellowship.

Invited Presentations (>100 invited presentations over the past 10 years)

- 2017
  Vienna, Austria. Credit Suisse Thought Leadership Forum
  Davos, Switzerland. World Economic Forum
  Aegina, Greece. Summer school on Social Cognition
  Beirut, Lebanon. Department of Philosophy. American University of Beirut
- 2016
  Cape Town, South Africa. South African Neuroscience Society Symposium
  Berlin, Germany. Max Planck Center for Adaptive Rationality
  Tehran, Iran. International Brain Research Organization meeting
  Shiraz, Iran. Shiraz University of Medical Sciences Seminar series
  Aegina, Greece. Summer school on Social Cognition
  Rusutsu, Japan. Winter Symposium on Brain and Mind.
- 2015
  Rio de Janeiro, Brazil. Conference on Financial Education and Investor behaviour
  Copenhagen, Denmark. Principia meeting on Social Neuroscience
  Nijmegen, Netherlands. International Symposium on Prediction and Decision Making
  Fondacion Les Treilles, Frace Fondacion. Subjective Confidence Meeting.
  Berlin Free University. Germany. Departmental Seminar Series
  Tehran, Iran. Iranian Cognitive Science Center meeting
  Aarhus, Denmak. Interacting Mind Center Meeting
  Cawley Manor, UK. Bridges Project Retreat. Open Society Foundation
  London, UK. NESTA. The roots of collective intelligence (http://bit.ly/2iAtAxD)
- 2014
  Estoril, Portugal. Champalimaud Centre. “One, Two, Many Brains” meeting
  Budapest, Hungary. Central European University Seminar series on Social Cognition
  London, UK. British Council meeting to promote Interaction between Iran and UK universities
  London, UK. Royal Institution Lecture series on Psychology
  Tehran, Iran. Tehran University of Medical Sciences Seminar on Consciousness
  Trondheim, Norway. Seminar Series on Cognitive Neuroscience
  Siena, Italy. Departmental Seminar Series on Cognitive sciences
- 2013
  Bristol, UK. Departmental Seminar Series on Human Decision Making
  Paris, France. Ecole Normal Superior Workshop on Social Cognition
  Okazaki, Japan. Symposium on Metacognition
  Bremen, Germany. European Conference on Visual Perception Symposium on Preconscious processing
  Copenhagen, Denmark. DTU Computer Science departmental seminar
  London, UK. BNA Festival Exploring and Celebrating Neuroscience

References

1 Sharot, T., Korn, C. W. & Dolan, R. J. How unrealistic optimism is maintained in the face of reality. *Nature neuroscience* **14**, 1475-U1156, doi:10.1038/nn.2949 (2011).

2 Navajas, J., Bahrami, B. & Latham, P. E. Post-decisional accounts of biases in confidence. *Curr Opin Behav Sci* **11**, 55-60, doi:10.1016/j.cobeha.2016.05.005 (2016).

3 Kruger, J. & Dunning, D. Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *J Pers Soc Psychol* **77**, 1121-1134, doi:Doi 10.1037//0022-3514.77.6.1121 (1999).

4 Mahmoodi, A. *et al.* Equality bias impairs collective decision-making across cultures. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 3835-3840, doi:10.1073/pnas.1421692112 (2015).

5 Bahrami, B. *et al.* What failure in collective decision-making tells us about metacognition. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **367**, 1350-1365, doi:10.1098/rstb.2011.0420 (2012).

6 Bang, D. & Frith, C. D. Making better decisions in groups. *Roy Soc Open Sci* **4**, doi:ARTN 170193 10.1098/rsos.170193 (2017).

7 Danileiko, I. & Lee, M. D. A Model-Based Approach to the Wisdom of the Crowd in Category Learning. *Cognitive Sci*, n/a-n/a, doi:10.1111/cogs.12561.

8 Aitchison, L., Bang, D., Bahrami, B. & Latham, P. E. Doubly Bayesian Analysis of Confidence in Perceptual Decision-Making. *PLoS computational biology* **11**, e1004519, doi:10.1371/journal.pcbi.1004519 (2015).

9 Sanders, J. I., Hangya, B. & Kepecs, A. Signatures of a Statistical Computation in the Human Sense of Confidence. *Neuron* **90**, 499-506, doi:10.1016/j.neuron.2016.03.025 (2016).

10 Ais, J., Zylberberg, A., Barttfeld, P. & Sigman, M. Individual consistency in the accuracy and distribution of confidence judgments. *Cognition* **146**, 377-386, doi:10.1016/j.cognition.2015.10.006 (2016).

11 Broihanne, M. H., Merli, M. & Roger, P. Overconfidence, risk perception and the risk-taking behavior of finance professionals. *Financ Res Lett* **11**, 64-73, doi:10.1016/j.frl.2013.11.002 (2014).

12 Nitzan, S., Paroush, J. *Collective Decision Making: An Economic Outlook.* (Cambridge University Press, 1985).

13 Yaniv, I. The benefit of additional opinions. *Curr Dir Psychol Sci* **13**, 75-78 (2004).

14 Hertz, U., Romand-Monnier, M., Kyriakopoulou, K. & Bahrami, B. Social Influence Protects Collective Decision Making From Equality Bias. *J Exp Psychol Human* **42**, 164-172, doi:10.1037/xhp0000145 (2016).

15 Bang, D. *et al.* Confidence matching in group decision-making. *Nature Hum Beh* **1**, 0117, doi:10.1038/s41562-017-0117 (2017).

16 Forsyth, D. R., Zyzniewski, L. E. & Giammanco, C. A. Responsibility Diffusion in Cooperative Collectives. *Personality and Social Psychology Bulletin* **28**, 54-65, doi:10.1177/0146167202281005 (2002).

17 Bayarri, M. J. & Degroot, M. H. Optimal Reporting of Predictions. *J Am Stat Assoc* **84**, 214-222, doi:Doi 10.2307/2289866 (1989).

18 Hertz, U. *et al.* Neural computations underpinning the strategic management of influence in advice giving. *Nat Commun* **8**, 2191, doi:10.1038/s41467-017-02314-5 (2017).

19 Tetlock, P. E. *Expert political judgment : how good is it? How can we know? : with a new preface by the author*. New edition. edn.

20 Campbell-Meiklejohn, D. K. *et al.* Structure of orbitofrontal cortex predicts social influence. *Current Biology* **22**, R123-R124, doi:DOI 10.1016/j.cub.2012.01.012 (2012).

21 Edelson, M., Sharot, T., Dolan, R. J. & Dudai, Y. Following the Crowd: Brain Substrates of Long-Term Memory Conformity. *Science* **333**, 108-111, doi:10.1126/science.1203557 (2011).

22 Izuma, K. & Adolphs, R. Social manipulation of preference in the human brain. *Neuron* **78**, 563-573, doi:10.1016/j.neuron.2013.03.023 (2013).

23 Salganik, M. J., Dodds, P. S. & Watts, D. J. Experimental study of inequality and unpredictability in an artificial cultural market. *Science* **311**, 854-856, doi:10.1126/science.1121066 (2006).

24 Condorcet, M. *Essai sur l'application de l'analyse á la probabilité des décisions rendues á la pluralité des voix.*, (de l'Impr. Royale, 1785).

25 Jowett, B. *The republic of Plato*. (Clarendon Press, 1888).

26 Galton, F. Vox populi. *Nature* **75**, 450-451, doi:Doi 10.1038/075450a0 (1907).

27 Surowiecki, J. *The Wisdom of Crowds*. (Doubleday; Anchor, 2005).

28 Forsythe, R., Nelson, F., Neumann, G. R. & Wright, J. Anatomy of an Experimental Political Stock-Market. *Am Econ Rev* **82**, 1142-1161 (1992).

29 Keller, A. *et al.* Predicting human olfactory perception from chemical features of odor molecules. *Science* **355**, 820-+, doi:10.1126/science.aal2014 (2017).

30 Kurvers, R. H. J. M. *et al.* Boosting medical diagnostics by pooling independent judgments. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 8777-8782, doi:10.1073/pnas.1601827113 (2016).

31 Mackay, C. *Extraordinary Popular Delusions and the Madness of Crowds*. 4th edn, (Wordsworth Editions Limited, 1841).

32 Pescetelli, N., Rees, G. & Bahrami, B. The perceptual and social components of metacognition. *J Exp Psychol Gen* **145**, 949-965, doi:10.1037/xge0000180 (2016).

33 Hastie, R. & Kameda, T. The robust beauty of majority rules in group decisions. *Psychological review* **112**, 494-508, doi:10.1037/0033-295X.112.2.494 (2005).

34 Raafat, R. M., Chater, N. & Frith, C. Herding in humans. *Trends in cognitive sciences* **13**, 420-428, doi:10.1016/j.tics.2009.08.002 (2009).

35 Tversky, A. & Kahneman, D. Judgment under Uncertainty - Heuristics and Biases. *Science* **185**, 1124-1131, doi:DOI 10.1126/science.185.4157.1124 (1974).

36 Chari, V. V. & Kehoe, P. J. Financial crises as herds: overturning the critiques. *J Econ Theory* **119**, 128-150, doi:10.1016/S0022-0531(03)00225-4 (2004).

37 Berkowitz, L. & Macaulay, J. Contagion of Criminal Violence. *Sociometry* **34**, 238-+, doi:Doi 10.2307/2786414 (1971).

38 Navajas, J. *et al.* The idiosyncratic nature of confidence. *Nature Human Behaviour*, doi:10.1038/s41562-017-0215-1 (2017).

39 Bahrami, B. *et al.* Optimally interacting minds. *Science* **329**, 1081-1085, doi:10.1126/science.1185718 (2010).

40 Regan, D. A High Frequency Mechanism Which Underlies Visual Evoked Potentials. *Electroen Clin Neuro* **25**, 231-&, doi:Doi 10.1016/0013-4694(68)90020-5 (1968).

41 Sajda, P., Philiastides, M. G. & Parra, L. C. Single-Trial Analysis of Neuroimaging Data: Inferring Neural Networks Underlying Perceptual Decision-Making in the Human Brain. *IEEE Reviews in Biomedical Engineering* **2**, 97-109, doi:10.1109/RBME.2009.2034535 (2009).

42 Zizlsperger, L., Sauvigny, T., Händel, B. & Haarmeier, T. Cortical representations of confidence in a visual perceptual decision. *Nat Commun* **5**, 3940, doi:10.1038/ncomms4940 (2014).

43 Gherman, S. & Philiastides, M. G. Human VMPFC encodes early signatures of confidence in perceptual decisions. *bioRxiv*, doi:10.1101/224337 (2017).

44 Weigelt, K. & Camerer, C. Reputation and Corporate-Strategy - a Review of Recent Theory and Applications. *Strategic Manage J* **9**, 443-454, doi:DOI 10.1002/smj.4250090505 (1988).

45 Frith, C. D. & Frith, U. Interacting minds--a biological basis. *Science* **286**, 1692-1695 (1999).

46 Chiu, S. *et al.* Anterior cingulate volume in pediatric bipolar disorder and autism. *J Affect Disord* **105**, 93-99, doi:10.1016/j.jad.2007.04.019 (2008).

47 De Martino, B., O'Doherty, J. P., Ray, D., Bossaerts, P. & Camerer, C. In the mind of the market: theory of mind biases value computation during financial bubbles. *Neuron* **79**, 1222-1231, doi:10.1016/j.neuron.2013.07.003 (2013).

48 Navajas, J., Niella, T., Garbulsky, G., Bahrami, B. & Sigman, M. Aggregated knowledge from a small number of debates outperforms the wisdom of large crowds. *Nature Human Behaviour*, doi:10.1038/s41562-017-0273-4 (2018).

49 Lorenz, J., Rauhut, H., Schweitzer, F. & Helbing, D. How social influence can undermine the wisdom of crowd effect. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 9020-9025, doi:10.1073/pnas.1008636108 (2011).