European Research Council

ERC Starting Grant
Research proposal (Part B section 1 (B1))[1]
(to be evaluated in Step 1)

Proposal Full Title: Neurobiological basis of collective decision making in the human brain

PROPOSAL ACRONYM: NeuroCoDec

Cover Page:

- Bahador Bahrami
- University College London
- Neurobiological basis of collective decision making in the human brain
- Neurobiology of cooperation
- 60 months

Proposal summary (half page)

We continually interact with each other and share information to make decisions together as friends, families, committees, juries, interest groups and institutions. The question of how collective decisions are made dates back many centuries and has been vigorously studied in social psychology and political economics, but the biological basis of collective decision making in human brain is almost entirely unknown. This burgeoning question is arguably (http://bit.ly/g3nvnx) at the heart of human society's current and urgent need to communicate effectively and find better ways of arriving at global collective decisions. My research has provided new and potentially important computational models of collective decision-making based on empirical data from visual psychophysics. I now propose to characterise the neurobiological underpinnings of joint decisions by combining recent advances in economics and social cognitive neuroscience in a novel interdisciplinary research program that will ask four questions: (1) How do we learn to make better collective decisions? (2) What are the brain mechanisms underlying the different psychological components of collective decision making? (3) What makes some people better or worse at collective decision making? Do the brains of effective collaborators complement or copy each other? (4) What is the role of the neuro-modulatory hormones oxytocin and testosterone in collective decision making? I plan to develop a formal theoretical model of collective decision making (Q1) and use converging evidence from complementary methodologies e.g. behavioural experiments & functional magnetic resonance imaging (fMRI) (Q2), structural brain imaging (Q3) and neuro-pharmacological investigations (Q4) to test and advance the model. This project will shed light on the neural mechanisms that combine the representation of our opinions with those of others. My findings could also help understand what physiological events underlie our ability to learn from experience to contribute effectively to group efforts and also disclose the biological basis that makes some of us better and some worse at working in groups . Finally, the results may provide a clearer picture of how hormonal interactions in the brain strike a delicate balance between trusting ourselves and accepting the opinion of others.

[1] Instructions for completing Part B section 1 (B1) can be found in the Guide for Applicants for the Starting Grant 2012 Call.

Section 1: *The Principal Investigator* 1(a) Scientific Leadership Potential (max 1 page)
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 2002 and still a medical student, I was invited to visit the Vision Lab at University of British Columbia in Vancouver, Canada. Hosted by Prof. Vince Di Lollo, I spent two months in UBC as a visiting fellow and worked on a project on visual selective attention. My first, single-author journal paper was published in 2003. My MD thesis investigated resting state brain activity in bipolar mood disorder and was published in 2005.

Starting in 2004 in London, my PhD focused on the distinction between consciousness and selective attention. I challenged the traditional notion in psychology that what is unconscious is also automatic, effortless and unaffected by attention. My findings were covered extensively by the media e.g. http://bbc.in/5YLj30. I finished my PhD with 4 publications as corresponding author and contributed to 2 other papers as collaborator (publications no. 18, 22) and supervised 2 students in their projects (no. 12, 15, 16).

I then joined the Interacting Mind (IM) Project in Aarhus University, Denmark in 2009 as a postdoctoral fellow and over the course of the next 2 years I set up a unique collaboration between the Awareness Lab in ICN UCL, the IM Project in Aarhus and Gatsby Computational Neuroscience Unit in UCL. I combined my expertise in visual psychophysics and signal detection theory with the IM project's focus on social cognition and developed a novel theoretical and empirical framework for investigating interpersonal communication. The first result of this multi-centre partnership was a paper published in *Science* in 2010 [1]. The paper revisited two contrasting maxims: "Two heads are better than one" and "Too many cooks spoil the broth" and asked: under what condition(s) is cooperation beneficial or damaging? To answer this question, for the first time, I tested *pairs* of observers in a visual sensory decision making task and measured their respective individual and joint decision making sensitivity. Based on my empirical observations, I developed a computational model of joint decision-making by *confidence sharing* which made a number of counter-intuitive predictions that were empirically supported by 3 further experiments also reported in the same paper. Most importantly, the model and the results identified the quantitative limit for transition from collective benefit ("two heads …") to failure ("too many cooks …"). The paper's potential for opening new avenues for research was immediately recognized in peer commentaries (Ernst, *Science* 329:1022-3; Dehaene & Changeux, *Neuron* 70:200-7; Skoyles, *Science* 330:1477).

I continued to explore and investigate the psychological characteristics of collective decision making and relate my theoretical approach and empirical findings to the rich literature on group dynamics in social psychology. In 2010 I was awarded a British Academy Postdoctoral Fellowship to study numerous psychological aspects of collective decision making by addressing the role of language [2], knowledge of outcomes [3] and mode of communication[4]. I have extended my laboratory model to more naturalistic tasks such as enumeration [5] and collective wagering & have studied the impact of personality traits on collective decisions. I have initiated collaborations with Beijing University, China and Tehran University, Iran and conducted cross-cultural comparative studies comparing Denmark, UK, Iran and China.

Now, with a firm grasp of the psychological basis of social collective decision-making, I seek to investigate its neurobiological underpinnings using the experience in techniques such as fMRI that I have gained during my training. To fulfil this vision, I am applying for the ERC 2012 StG as a STARTER.

In preparation for this project, I have established expertise in model-based functional imaging (see publications 5 & 14) and have supervised an fMRI project (see publications 13). As a result, I am now confident that I can use brain imaging (fMRI) effectively to investigate the neural substrates of collective decision making. I have recently contributed to a series of studies on the biological basis of individual differences in cognition and their relationship to brain structure. This line of work puts me in a strong position to address neuro-structural basis of variability in cooperative success. Finally, I have investigated the role of the human sex hormone testosterone in collective decision making [6]. The results of this project (see Synopsis) suggest significant potential for this paradigm [1] in disclosing the biological aspects of social interaction, as employed in this application. Importantly, this project has helped me familiarize with the ethical and administrative requirements of this type of scientific approach.

1(b) Curriculum Vitae (max 2 pages)

Current status

Since Sept 2010 **British Academy Postdoctoral Research Fellow**

- Institute of Cognitive Neuroscience (ICN)
- Project title: Are two heads better than one? The cognitive neuroscience of collective decision making.

Mentor: Prof. Geraint Rees

Funding ID

2010 – 2013 **British Academy Post-Doctoral Research Fellowship**
Total grant amount: £247,733.00

2004 – 2007 Graduate School Research Scholarship (GSRS), UK
Fund value: £60,300.00

2004 – 2007 Overseas Research Scholarship (ORS), UK
Fund value: £30,000.00

Previous position

2008 – 2010 **Postdoctoral Research Fellow jointly at**

- Neils Bohr Interacting Mind Project, Aarhus University, Aarhus, Denmark
  - Supervisors: Prof. Chris Frith and Prof. Andreas Roepstorff (focus: Social Cognition)
- Institute of Cognitive Neuroscience (ICN) and Wellcome Dept of Imaging Neuroscience, UCL, London UK
  - Supervisor: Prof. Geraint Rees (focus: Awareness)

Education

2004 – 2008 **PhD, University College London, Institute of Cognitive Neuroscience & Psychology Department (Fully funded by UK-Universities Graduate School Research (GSRS) & Overseas Research (ORS) Scholarships**
Supervisors: Prof. Vincent Walsh, Prof. Nilli Lavie
Title of Dissertation: Attention without awareness.

1995 – 2003 **Medical Doctorate, Tehran University of Medical Sciences, Tehran, IRAN**

Previous Research

2001 – 2003 **Medical Student Researcher**
Electroencephalography Clinic, Roozbeh Psychiatry Hospital, Tehran University of Medical Sciences, Tehran, IRAN

2002 **Invited Visiting Researcher**
Vision Lab, Psychology Department, University of British Columbia, Vancouver, Canada; Host: Prof. Vincent Di Lollo

1997 – 1999 **Volunteer Research Assistant**
Laboratory of Excimer Lasers, Tehran Polytechnique, Tehran, IRAN

Teaching Experience

- Annual summer school in Experimental Methods in Psychology/Neuroscience at School of Cognitive Sciences, Iran (July 2008-2011). Topics
  1) Visual Perception
  2) Quantitative Methods
  3) Signal Detection Theory
- Matlab for Cognitive Neuroscience at UCL/ICN (2008-2011).
- MSc in Cognitive Neuroscience at UCL (course coordinator: Leun Otten); Topic: Collective Decision Making

Co-Organizer **2nd IMPRS NeuroCom Summer School, London UK, July 6 – 8 2011**
Principle organizer: Prof. Patrick Haggard
http://imprs-neurocom.mpg.de/summerschool/index.html

Supervised BSc, MSc and PhD student

During my time as a PhD and Post doc in the ICN, UCL as well as in Aarhus University, Denmark and Tehran University, IRAN, I have contributed to the training of other students and post docs. I have taken on numerous collaborative, teaching and mentoring duties and have been responsible for the success of several new projects. Below you will find a list of students for whom I have acted as the mentor.

1. Petra Vetter. PhD student in ICN - project theme: the role of attention in enumeration – see publications no. 13, & 16
2. Mia Y Dong. BSc student in ICN - project theme: distractibility in everyday life and brain structure – see publication no. 6
3. Karsten Olsen. MSc student in Aarhus University, Denmark - project theme: collective decision making – see publications no. 3, 12.
4. Dan Bang. BSc student in Aarhus University, Denmark - project theme: collective decision making – see publications no. 3
5. Julia Harris. PhD student on a rotation in ICN - project theme: the limits of unconscious processing – see publications no. 10
6. Eva Spolaore. MSc student in University of Padova, Italy - project theme: unconscious enumeration – see publications no. 15.
7. Silvia Pagano. MSc student in University of Padova, Italy - project theme: unconscious enumeration – see publications no. 15.
8. Kristian Sandberg. PhD student in Aarhus University, Denmark - project theme: neural correlates of awareness; see publication no 7 – and poster presented in Vision Science Society meeting 2011, Naples, FL. http://bit.ly/nnqiFX
9. Nahid Zokaei. BSc and PhD student in UCL, UK - project themes: Semantic and perceptual motion processing; see publication no. 17 – Working memory for motion, poster presented in Vision Science Society meeting 2011, Naples, FL. http://bit.ly/niaUsI
10. Nikos Konstantinou. PhD student in UCL, UK – project theme: working memory and attention. Poster presented in Vision Science Society meeting 2010 http://bit.ly/qaWOOs
11. Moritz Stolte. PhD student in UCL, UK – project theme: Attentional load and sensory precision. Poster presented in Vision Science Society meeting 2011. http://bit.ly/quJ7AK
12. Thomas Ditye. PhD student in UCL, UK – project theme: Perceptual learning. Poster presented in Vision Science Society meeting 2011. http://bit.ly/p4Qp33
13. Christos Sideras. MSc student at UCL, UK – project theme: Collective wagering. Thesis submitted in Sept 2011.
14. Payam Piray. MSc student at Tehran University, IRAN – project theme: The role of perceived complexity in risky decision making. Paper under review.
15. Spas Getov . MSc student in UCL, UK – project theme: Individual differences in unconscious face perception. Poster presented in Vision Science Society meeting 2011. http://bit.ly/n6lKZZ
16. Sara Ajina. MSc student in UCL, UK – project theme: Unconscious social evaluation of faces. Poster presented in Vision Science Society meeting 2011. http://bit.ly/n6lKZZ
17. Lorna Stewart. PhD student in UCL, UK – project theme: The role of personality traits in unconscious face perception. Poster presented in Vision Science Society meeting 2011. http://bit.ly/n6lKZZ

1(c) Early Achievement-Track-Record (max 2 pages)

Publications	**Citation**
*: papers without my PhD supervisors	
†: corresponding author	
‡: project supervisor	
1. * Zokaie, N., **Bahrami, B.**, Gorgoraptis, N., Bays, P., and Husain, M. Precision of working memory for visual motion and transparent motion surfaces. *Journal of Vision* (accepted)	0
2. * Kanai, R., **Bahrami,B.**, Roylance, R., Rees, G. (2011) Online social network size is reflected in human brain structure. *Proceedings of the Royal Society B. DOI:* 10.1098/rspb.2011.1959	0
3. *† **Bahrami, B.,** Olsen, K., Bang, D., Roepstorff, A., Rees., G. & Frith, C. (in press) Together, slowly but surely: the role of social interaction and feedback in the build-up of benefit in collective decision-making. *JEP:HPP*.	0
4. * Meteyard, L., Rodriguez, S., **Bahrami, B.,** Vigliocco, G. (in press). Beyond the current state of the art: a review of embodiment and the neuroscience of semantics. *Cortex*. DOI: 10.1016/j.cortex.2010.11.002.	0
5. Cohen Kadosh, R., **Bahrami, B**, Walsh, V., Butterworth, B., Popescu, T., Price, C.J. (2011). Specialization in the human brain: the case of numbers. *Front. Hum. Neurosci.* **5**:62. *doi*: 10.3389/fnhum.2011.00062.	0
6. * Kanai, R., Dong, M., **Bahrami, B.**, Rees, G. (2011). Distractibility in daily life is reflected in the structure and function of human parietal cortex. *Journal of Neuroscience.* 31(18): 6620-6626.	0
7. * Sandberg, K., **Bahrami, B.**, Lindeløv, J., Overgaard, M., & Rees, G. (2011). The impact of stimulus complexity and frequency swapping on stabilization of binocular rivalry. *Journal of Vision* 11(2):6,1-10. doi: 10.1167/11/2/6	0
8. * Kanai, R., Carmel, D., **Bahrami, B.**, Rees, G. (2011) Structural and functional fractionation of right superior parietal cortex in bistable perception *Current Biology* 21(3):R106-R107.	1
9. *† **Bahrami, B.,** Frith, C. (2011) Interacting Minds: A Framework for Combining Process- and Accuracy-Oriented Social Cognitive Research. *Psychological Inquiry*. 22(3): 183-6.	0
10. *‡ Harris, J. J., Schwarzkopf, D. S., Song, C., **Bahrami, B**., Rees, G. (2011) Contextual illusions reveal the limit of unconscious visual processing. *Psychological Science*. 22(3):399-405.	0
11. * Kanai, R., **Bahrami, B.**, Rees, G. (2010) Human parietal cortex structure predicts the individual differences in perceptual rivalry. *Current Biology.* 20: 1626-30.	9
12. *† **Bahrami, B.**, Olsen, K., Latham, P., Roepstorff, A., Rees, G., Frith, C. (2010). Optimally interacting minds. *Science*. 329: 1081-1085.	4
13. *‡ Vetter, P., Butterworth, B. and **Bahrami, B.** (2010). A candidate for the attentional bottleneck: Set-size Specific Modulation of Right TPJ during Attentive Enumeration. *Journal of Cognitive Neuroscience* 23:1-9	1
14. Bueti, D., **Bahrami, B.**, Walsh, V. and Rees, G. (2010) Representation of temporal expectations in human primary visual cortex. *Journal of Neuroscience* 30(12):4343-4352.	7
15. *† **Bahrami, B.**, Vetter, P., Spolaore, E., Pagano, S., Butterworth, B., Rees, G. (2010) Unconscious numerical priming despite interocular suppression *Psychological Science* 21(2): 224-233.	4
16. *‡ Vetter, P., Butterworth, B., and **Bahrami, B.** (2008) Modulating attentional load affects numerosity estimation: Evidence against a pre-attentive subitizing mechanism. *PLoS ONE*.	8
17. *‡ Meteyard, L., Zokaei, N., **Bahrami, B.** & Vigliocco, G. (2008) Visual motion interferes with lexical decision on motion words. *Current Biology* 18(17), R732-R733.	5
18. † **Bahrami, B.**, Carmel, D., Walsh, V., Rees, G., Lavie, N. (2008). Spatial attention can modulate unconscious orientation processing. *Perception*. 37(10):1520-8.	7
19. Bueti,D., **Bahrami,B.**, Walsh,V. (2008) Sensory and association cortex in time perception. *Journal of Cognitive Neuroscience* 20: 1054-1062.	30
20. *† Kaul, C., **Bahrami, B.** (2008). Subjective experience of motion or attentional selection of a categorical decision. *Journal of Neuroscience* 28(16), 4110-4112.	6
21. † **Bahrami, B.**, Carmel, D., Walsh, V., Rees, G., & Lavie, N. (2008). Unconscious	4

orientation processing depends on perceptual load. *Journal of Vision*, *8*(3):12, 1-10, http://journalofvision.org/8/3/12/, doi:10.1167/8.3.12.

22. † **Bahrami, B.** Lavie, N. and Rees, G. (2007) Attentional load modulates responses of human primary visual cortex to invisible stimuli. *Current Biology* 17, 509-513. 35
23. * Meteyard,L., **Bahrami,B.**, Vigliocco,G. (2007). Motion detection and motion verbs: Language affects low-level visual perception. *Psychological Science* 18(11), 1007-1013. 23
24. *† **Bahrami, B.**, Seyedsadjadi, R., Babadi, B., Noroozian, M. (2005). Brain complexity increases in mania. *NeuroReport* 16, 187-191. 14
25. *† **Bahrami, B.** (2003). Object property encoding and change blindness in multiple object tracking. *Visual Cognition* 10(8), 949-963. 24

Prizes and Awards

2004-2007	Graduate School Research Scholarship (GSRS), UK – fund value: £60,300.00
2004-2007	Overseas Research Scholarship (ORS), UK – fund value: £30,0 00.00
2005-2006	Sully Scholarship, UK
2005	Oliver Braddick Prize in Psychology
2007	Vision Science Society Student Travel Award
2007	Brain Journal Student Travel Award
2010-2013	British Academy Post-Doctoral Research Fellowship – fund value: £247,733.00

Invited Conference Presentation

May, 2007	Vision Science Society Meeting, Sarasota, FL
May, 2011	Vision Science Society meeting, Naples, FL
June, 2011	Royal Society UK-German Frontiers of Science Meeting, Symposium on Consciousness

Invited presentations at International Advanced Schools

Sept, 2009	Marseille-London Summer School in Cognitive Neuroscience
July, 2008, Aug, 2009 and July 2010	Summer School in Cognitive Neuroscience Methods, School of Cognitive Sciences, Tehran, IRAN
July, 2011	Summer School in Neuro-Imaging Methods, School of Cognitive Sciences, Tehran, IRAN

Section 1d: *Extended Synopsis of the project proposal* (max 5 pages)

a. State-of-the-art: From 'Social' to 'Biological' in collective decision making

I seek to understand the neurobiological basis of collective decision making in humans. We continually interact with each other and share information to make decisions together: as friends, families, committees, soldiers, politicians, juries, interest groups and institutions (banks, consumers and citizens in democracies). However, this process is by no means trivial, takes more than simply sharing opinions between people and is far from guaranteed to succeed. Indeed, groups fail to do better than their best individuals alarmingly often and contradictory maxims such as "Two heads are better than one" and "Too many cooks spoil the broth" nicely illustrate the thin line between failure and success in cooperation.

The question of how collective decisions are made dates back many centuries [7] and has been widely studied in social psychology [8]. Many factors promote or impair joint decisions [9]. However, the biological basis of collective decision making in the human brain is almost entirely unknown. Recent advances in social cognitive neuroscience have now equipped us with conceptual and technological tools that enable us to address this unresolved question which is arguably at the heart of human society's current and urgent need to communicate effectively and find better ways of arriving at global collective decisions (http://bit.ly/g3nvnx). Grounded in the history of social psychology [8] and political economics [10], my recent research has now formalized some of the key psychological components of collective decision making. The stage is now set to address its neurobiological underpinning. To do this, here I propose an ambitious but realistic interdisciplinary research program focusing on 4 key questions:

1) **How do we learn to make better collective decisions?**
2) **What are the functional brain mechanisms underlying the different psychological components of collective decision making?**
3) **What makes some people better or worse at collective decision making? Do the brains of effective collaborators complement or copy each other?**
4) **What is the role of the neuro-modulatory hormones oxytocin and testosterone in collective decision making?**

This proposal has 4 sections. Each explains a problem and the methodology (including sample size & feasibility considerations) that will be used to address it. The common theme connecting the questions is to use the converging evidence from complementary fields to develop and test a formal theoretical model of collective decision making.

b. Methodology: A model system for collective decision-making. I have recently developed a laboratory model for collective decision making [1,3,11]. In this paradigm (Fig 1), pairs of participants (dyad) together view a brief visual display where a faint target (contrast oddball; Fig1) occurs in either the 1st or 2nd viewing interval. Initially, each participant independently chooses the interval they think contains the oddball, without consulting the other. Individual decisions are then shared, and if there is disagreement, participants negotiate until a joint decision is reached. The correct choice is then declared. Individual and dyadic sensitivity and bias and the contribution of each person to the collective decision are quantified [1]. The ratio of dyad sensitivity to that of the more sensitive participant quantifies the benefit collective decision making. An ideal computational Bayesian model delineates the upper bound of performance achievable through optimal combination of individual opinions. The ratio of dyad sensitivity to this optimal bound determines how well the dyad has fulfilled its potential.

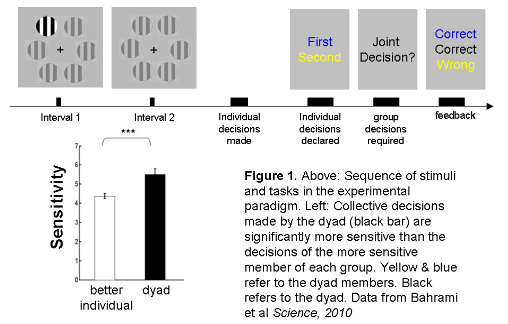

Figure 1. Above: Sequence of stimuli and tasks in the experimental paradigm. Left: Collective decisions made by the dyad (black bar) are significantly more sensitive than the decisions of the more sensitive member of each group. Yellow & blue refer to the dyad members. Black refers to the dyad. Data from Bahrami et al *Science, 2010*

Social interaction significantly improves dyadic perceptual sensitivity (Fig 1) and this benefit is consistently observed across various configurations of the paradigm such as enumeration [5] and joint wagering, with verbal [2] and nonverbal [4] communication modes and even across cultures (tested in UK, Denmark, Iran and China). I have recently compiled a battery of psychological assessment that provides a social personality profile strongly predictive of collective performance. These findings suggest that this model system is not restricted to visual contrast detection but reliably generalises to practically relevant forms of joint decision encountered in everyday life. Moreover, the close concordance between the results and the model's predictions suggest that the model captures key features of collective decision-making. This paradigm and the underlying model provide the theoretical and psychological foundation for investigating the neurobiological basis of collective decision-making described here.

1) **How do we learn to make better collective decisions?**

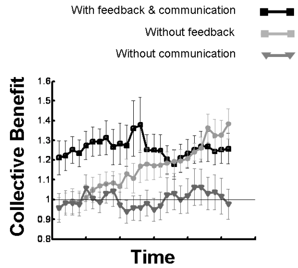

Figure 2. Social learning in the absence of feedback. Collective benefit is plotted against time. Values above the y=1 horizontal line indicate that dyad did better than individuals. Communicating dyads with feedback (black symbols) showed a stable and robust collective benefit across time. Non-communicating dyads did not show any benefit. Without feedback (light grey), collective benefit was gradually achieved as individuals learned to collaborate. Data from Bahrami et al *JEP:HPP 2011*.

Effective group decision-making requires much coordination and practice. However, surprisingly little is known about how people learn to contribute to joint decisions. Moreover, in real life, many joint decisions must be made without access to immediate outcomes, often because they are too far in the future e.g. parenting decisions. I recently showed that social interaction enables the build-up of collective benefit even when the decision outcomes are unknown [3]. The paradigm depicted in Fig 1 (above) was used except the dyads never knew who was right or wrong. Dyads did not initially achieve any collective benefit (Fig 2, light grey) but improved over time and by the end equalled the dyads with outcomes (Fig 2, black). This finding is inconsistent with my own previous model [1] which naively assumed that collective decision making is a stable process. Current computational models of social learning are based on principles of associative reinforcement learning (RL) [12-14] and invariably depend on outcomes for updating the learning process. They cannot, therefore, account for collective decision-making without feedback [3] (fig 2) either. Moreover, these models[12-14] are concerned with inferring hidden intentions in the face of conflict of interest whereas, collective decision making is a problem of information integration that persists even without conflict of interest (e.g. how to invest the family savings). I will therefore develop a new model to account for the build-up of collective decision making by improving information integration across individuals in the absence of conflict of interest, and will extend it to situations where outcome is inaccessible.

Description of the model. I propose that learning to make effective collective decisions requires 3 components. Group members should be able to (i) express their own confidence ($\boldsymbol{Z_{self}}$) accurately; (ii) estimate the others' confidence ($\boldsymbol{Z_{other}}$) and (iii) combine the two effectively using the right decision rule. I assume that the absolute values of $\boldsymbol{Z_{self}}$ and $\boldsymbol{Z_{other}}$ are monotonically related to the probability of correct response (1^{st} or 2^{nd} interval) but corrupted by neural and environmental noise. The sign of these values indicate the decisions (-1 = 1^{st}, +1 = 2^{nd} interval). Previously, I have proposed [1] that $Z \propto \Delta C \times S$ where $\Delta \boldsymbol{C}$ is oddball contrast and $\boldsymbol{S}$ is the observer's sensitivity. Consequently, to satisfy components (i) and (ii), members should learn $\boldsymbol{S_{self}}$ and $\boldsymbol{S_{other}}$. I define the decision rule by $d = f\left(Z_{self} + \beta Z_{other}\right)$ which gives the probability of confirming the self decision; here $\boldsymbol{\beta}$ is a weighting factor and $f(x) = 1/\left(1 + e^{-\omega x}\right)$ is the logistic sigmoid with $\boldsymbol{\omega}$ reflecting the noise in the dyadic decision. I have studied [1] a special case ($\boldsymbol{\beta}$=*1*) of this formulation under Gaussian assumptions. Component (iii) of social learning seeks to find a value of $\boldsymbol{\beta}$ that maximizes dyad accuracy. Following [13], I will start by formulating my model according to the principles of reinforcement learning. The model gives a trial-by-trial estimate of $\boldsymbol{S_{self}}$, $\boldsymbol{S_{other}}$ and $\boldsymbol{\beta}$ which are updated according to the Rescorla-Wagner rule. For example, after trial $\boldsymbol{t}$, $\boldsymbol{S_{self}}$ is updated by $S_{self}^{t+1} = S_{self}^{t} + \eta \delta_{self}^{t}$ in which δ is the confidence prediction error: $\delta_{self}^{t} = R_t - Z_{self}^{t}$ where $\boldsymbol{R_t}$ is the correct choice in trial $\boldsymbol{t}$. The model will then be extended to no-feedback cooperation and tested against empirical data (Fig 2, [3]) by replacing $\boldsymbol{R}$ with dyadic decision (i.e. accepting the joint decision as outcome).

Objectives and implications. Once fully developed, I will fit the model to a large database of behavioural experiments I have already conducted using the paradigm described above. This will allow me to identify the computational strategies employed by interacting humans learning to cooperate effectively. Moreover, the resulting model will be employed in the analysis of the fMRI data obtained in Experiments 2 & 3 of section 2. Successful extension of the model to no-feedback condition would suggests that a functional role of shared subjective experience i.e. awareness, may be to replace missing reinforcement when decision outcomes are not available, too complex to estimate or far in the future. Given the abundance of situations in everyday life where immediate outcomes are difficult, sometimes even impossible to determine, this idea may offer an ecologically relevant social role for conscious awareness.

Collaborators: One postdoctoral Fellow with a PhD in computational neuroscience/decision-making will be hired. I will collaborate with Prof. Peter Latham from the Gatsby Computational Neuroscience Unit, UCL.

2) What are the functional brain mechanisms underlying the various psychological components of collective decision making?

Recent research in neuroeconomics has identified the mechanisms involved in decision making in isolated individuals that assess and accumulate the evidence favouring different choice option [15], compare the likelihood and magnitude of rewards associated with them [16] and execute the action that implements the decision [17]. In parallel, research in social neuroscience has developed a comprehensive theoretical and empirical picture of the brain mechanisms that contribute to estimating others' intentions and goals [13,14,18]. There is now a growing consensus [12,19] that an urgent next step is to combine these two separate fields to understand the neural basis of human social behaviour. Collective decision making, as the paradigmatic case where social interaction and value-based decision making come together, is the natural setting for these fields to join. I propose three linked functional MRI experiments that combine these neuro-economic and socio-cognitive approaches to identify the neuronal substrates of joint decision-making.

Experiment 1: I hypothesize that (i) separate neural mechanisms code the self and other's confidence in the human brain; (ii) the weighted combination of self and other's confidence (according to the decision rule defined in question 1) is represented in the human brain. The paradigm will be optimized (Fig 3) for fMRI by a number of key changes following [4,18]. In every session, one recruited participant will be paired with an actor from the lab. The participant performs the task in a MRI scanner and the actor pretends to co-operate by doing the task outside the scanner. The actor's responses will be generated by the computer in a 2x2 factorial design (Fig 3) whereby the participant's (high and low) confidence trials are matched with the actor's high and low confidences. Using an adaptive staircase procedure [20], overall accuracy of participant and actor will be kept equal and stable at 75%. Throughout the experiment, the participant remains oblivious to the fact that s/he is collaborating with an actor whose responses are computer generated. In each trial, the participant first makes a private decision about target interval and indicates her/his confidence graphically by how far s/he drags a marker to the left (i.e. 1st interval) or right (2nd interval) of the centre (Fig 3). The actor's decision and confidence will then be declared to participant together with her/his own. She will then make a joint decision for the dyad. Finally feedback will be given. I will use Statistical Parametric Mapping techniques implemented in SPM8 to address my hypotheses. I predict that (i) main effect of Self (see 2x2 design in Fig 3) will identify the right rostral medial [20] and left dorsolateral [21] prefrontal cortex & (ii) main effect of Other will identify the anterior [18] and middle [22] cingulate cortex. (iii) A conjunction analysis will test if the neural representation of self and others are shared. (iv) The neural correlate of the combined confidences (i.e. decision varibale) may be the ventromedial prefrontal cortex (VMPFC) [17].

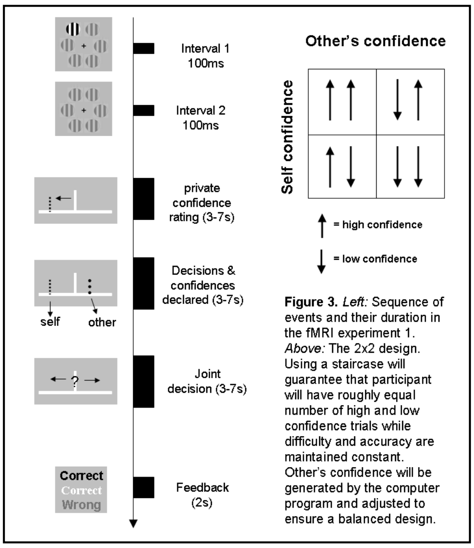

Figure 3. *Left:* Sequence of events and their duration in the fMRI experiment 1. *Above:* The 2x2 design. Using a staircase will guarantee that participant will have roughly equal number of high and low confidence trials while difficulty and accuracy are maintained constant. Other's confidence will be generated by the computer program and adjusted to ensure a balanced design.

Experiment 2: Using the paradigm of experiment 1 and the social learning model (section 1), I will now identify the neuronal correlates of social learning in collective decision making. Critically, here the computer will generate decisions for the actor so that her accuracy will fluctuate in blocks ~20 trials, sometimes more accurate than the participant (who is fixed at ~75%) and sometimes less. Each participant should therefore keep an ongoing track of her partner's accuracy and take her fluctuating reliability into account when making joint decisions. I will then apply the learning model to the behavioural data to infer trial-by-trial estimates of learned self- and other-reliability and the most-likely applied decision rule. I will fit these parameters to the whole-brain fMRI data using model-based fMRI analysis to identify the neuronal correlates of the learning process. Questions are: (i) which brain areas are involved in updating the participant's estimate of self reliability? Prediction: dorsolateral prefrontal cortex [17]. (ii) Which brain area is involved in updating the partner's reliability? Prediction: anterior cingulate cortex [18] (iii) Does the brain update its decision rule (presumably represented in VMPFC [17]) for how to combine self and other's confidence? (iv) Finally, I will test if any relevant brain areas (identified in experiment 1) contribute to social learning in a manner inconsistent with the predictions of the social collective learning model.

Experiment 3: I will study joint decision making in the absence of feedback. My behavioural preliminary data (Fig 2) [3] suggest that, with practice, social interaction can replace the outcome information. Experiment 3 will seek the neural basis of these effects. The experiment will be identical to experiment 1 except that the last stage of each trial (Fig 3; feedback) will be removed. In analysis, the data from Experiment 1-2 will be used to identify the neural correlates of self-, other- and collective error. I will then examine what happens in the brain in situations of disagreement in the absence of feedback e.g., a highly confident disagreeing partner will be interpreted as self- error in areas that tracked errors in experiments 1-2. Furthermore, the model-based approach (above) will be employed to test if the same brain areas that used outcome information in experiment 2 to update belief about reliability of self- and/or other would now employ the joint decision as a surrogate outcome to update beliefs.

Sample size and technical considerations: Principled power calculations are known to be difficult in neuroimaging mass-univariate analysis frameworks, so for experiments 1-3, I will adopt the approach now standard in the field of using previous similar studies to infer the sample size needed to reject the null hypothesis in one or more voxels. Based on [13,18] I will recruit N=25 adult healthy human volunteers for each experiment. Scanning will be performed in a Siemens Trio 3-Tesla MRI scanner. Whole brain Blood Oxygenation Level Dependent (BOLD) signal, as well as pulse rate and oximetry data will be recorded.

Collaborators: A postdoctoral research assistant with a PhD in human neuroimaging will be hired in 2013 for carrying out the project. The other RA (hired for section 1) will also contribute to the model-based fMRI analysis. I will collaborate with Prof. Geraint Rees from the UCL Institute of Cognitive Neuroscience.

3) What makes some people better and some others worse at collective decision making? Do the brains of effective collaborators complement or copy each other?

We are not all equally good at joint decisions. What is the biological source of this variability? We have recently shown that individual differences in a range of cognitive faculties spanning sensory perception [23], distractibility in every day life [24], social network size [25] & social influence by others [26] are reflected in the grey matter structure of localized regions in the human brain. I hypothesize that variations in local brain structure are predictive of individuals' ability to contribute to successful collective decisions.

Methods: In a large group of participants, structural MRI and diffusion tensor imaging (DTI) brain scans and social personality profiles will be obtained. The participants will also be paired to undertake the joint decision paradigm. Using the standard approach of Voxel Based Morphometry & Fractional Anisotropy Analysis in SPM8 (which I have used and taught repeatedly in the past 2 years), I will test whether local gray matter volume, & white matter connectivity correlate with the individuals' contribution to the correct joint decisions and social personality profiles. Based on our most recent work on social influence [26], I predict bilateral orbitofrontal cortex (OFC) will show such correlation.

My previous computational work and more recent linguistic analysis of collective decisions suggests that similarity in perceptual sensitivity [1] and convergence to a common, small set of linguistic markers to discuss confidence [2] are crucial predictors of collective success. Is similarity a more general predictor of success in joint decisions? In particular, what are the relational characteristics of successful collaborators' brains? Do the brains of effective collaborators complement or copy each other? To answer these questions, I will take the MRI images of each dyad and for each voxel I will calculate a distance index (DI) (i.e. the difference in grey matter volume between two participants) to construct a dyadic-distance brain (DDB). I will then test whether local variations in DI in the DDB correlate with dyadic collective benefit.

Sample size: Similar to section 2, I will adopt the approach standard in the field of using previous similar studies [23-25] and collect data from N=160.

Collaborators: Prof. Geraint Rees and Dr. Ryota Kanai at UCL Institute of Cognitive Neuroscience. Two MSc students from 2012 UCL cohort will be recruited to assist with the experiment.

4) What is the role of the neuro-modulatory hormones oxytocin and testosterone in collective decision making?

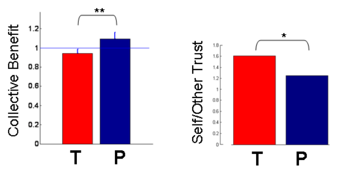

Figure 5. Left : Exogenous testosterone (red) reduced the collective benefit compared to placebo (blue). Right : testosterone also increased the participants' trust in their own decisions (right)

Research over the past decade has shown that understanding the role of neuro-hormonal modulators in interactive contexts is indispensible for social cognitive neuroscience [27]. I have recently discovered [6] (Fig 5) that testosterone causally and selectively disrupts individuals' ability to collaborate fruitfully. Specifically, testosterone impaired collaboration by increasing the individual's self-favouring choices indicating that testosterone reduces interpersonal trust. But striking a balance between self- and other-oriented decisions is essential for joint decisions. This suggests that an chemical agent must operate in the opposite direction to testosterone for the brain to maintain balance between self- and other-oriented decisions. I hypothesise that the nanopeptide oxytocin [28] which is involved in social bonding, emotional reaction, birth and lactation – is such a candidate.

Method: A randomised, double-blind, placebo-controlled within-subject design will examine the effect of oxytocin in joint decision making. Normal healthy adult volunteers will self-administer 20-60 IU nasal oxytocin or placebo. Nasal administration of oxytocin affects the processing of social (e.g. emotional faces) and non-social stimuli (e.g. fear conditioning) [28]. After 30-60 minutes, participants will take part in the collective decision task. I hypothesise that exogenous oxytocin will disturb the balance of self versus other trust in the opposite direction to testosterone and increase pro-social behaviour by promoting other-oriented arbitration in collective decisions. I hypothesise that the impact exogenous oxytocin on collective accuracy, however, will be similar to testosterone (Fig 5, left panel) because any factor that tips the balance of trust (in either direction) will harm joint decisions. The trust hypothesis will be formally investigated by applying the social learning model (section 1) to the data from the testosterone [6] and oxytocin treatments. The prediction is that these two treatments will have opposite effects on learning of self and other sensitivity and lead to biased decision rules that favour self (for testosterone) or the other (for oxytocin). The model will therefore serve as the link between this experiment and the fMRI data (section 2) to draw new hypotheses for the neural loci of action of testosterone and oxytocin in collective decision making.

Sample size and design: Based on my previous results [6], a sample size of N=40 participants (20 dyads) will provide adequate power for a repeated-measure study. In each session, both participants will receive the same treatment (placebo or drug). Order of treatments will be randomized across dyads. Participants and experimenters will be blind to the treatment in each session.

Collaborators: Prof. Geraint Rees and Prof. Chris Frith. Two MSc students from the 2016 UCL cohort will be recruited to assist with the experimental procedures.

Timetable: In 2012-13 (phase 1, Fig 6), the model will be developed and tested against previously collected data. Over the next two years (phase 2) the model will be extended analyze the fMRI experiments. The fMRI experiments will be carried out over the period of three years (Fig 6; and Table 1). Individual difference study will be carried out in 2012-13 (data collection: ~14 weeks. Analysis & write-ups: ~14 and ~20 weeks respectively. Oxytocin project will be carried out in 2016-17 (Fig 6). Preparation of the ethics application and approval: ~20 weeks. Data collection: ~10 weeks. Data analysis and write-ups: ~4 and 16 weeks respectively.

	Preparation of Paradigm for fMRI	Data collection	Analysis & write-up	Total
E1	12	10	24	46
E2	14	12	26	52
E3	8	10	20	38
All times are indicated in WEEKS.				

Table 1. Timtable of fMRI experiments.

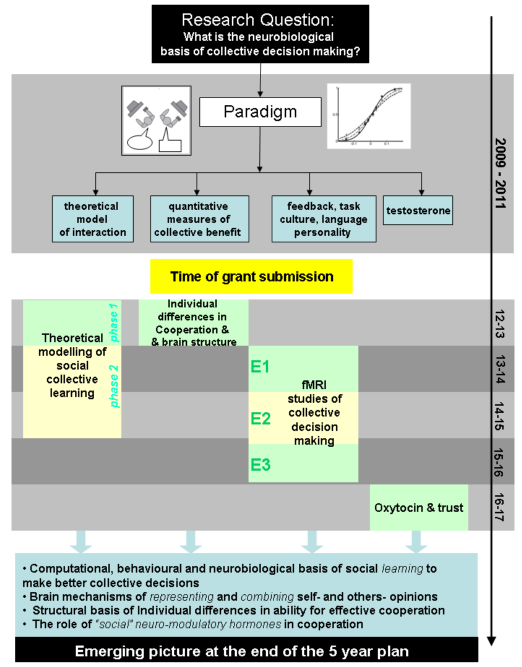

Figure 6. The proposal at a glance

References

1. Bahrami B., Olsen K., Latham P.E., Roepstorff A., Rees G., and Frith C.D. (2010). Optimally interacting minds. Science 329: 1081-1085.
2. Fusaroli R., Bahrami B., P, Olsen K., Roepstorff A., Frith C., and Tylen K. (2010). Unfolding cognitive coordination in a state of uncertainty. In Language as Social Coordination.
3. Bahrami B., Olsen K., Bang D., Roepstorff A., Rees G., and Frith C. (2011). Together, slowly but surely: The role of social interaction and feedback on the build-up of benefit in collective decision-making. J Exp. Psychol. Hum. Percept. Perform.
4. Bahrami B, Olsen K., Bang D., Roepstorff A, Rees G, and Frith C D (Under review). What failure in collective decision-making tells us about metacognition.
5. Bahrami B., Didino D., Frith C.D., Butterowth B., and Rees G (Under review). Collective enumeration.
6. Wright N., Bahrami B, Johnson E, DiMalta G, Rees G, Frith C.D., and Dolan R. (Under review). Testosterone disrupts human collaboration by increasing egocentric choices.
7. Condorcet M. (1785). Essai sur l'application de l'analyse á la probabilité des décisions rendues á la pluralité des voix. (Paris: de l'Impr. royale).
8. Kerr N.L. and Tindale R.S. (2004). Group performance and decision making. Annu. Rev. Psychol. 55: 623-655.
9. Hastie R. and Kameda T. (2005). The robust beauty of majority rules in group decisions. Psychol. Rev. 112: 494-508.
10. Nitzan S. and Paroush J. (1985). Collective decision making: an economic outlook. (Cambridge: Cambridge University Press).
11. Bahrami B. and Frith C.D. (2011). Interacting minds: a framework for combining process- and accuracy-oriented social cognitive research. Psychological Inquiry 22: 183-186.
12. Behrens T.E., Hunt L.T., and Rushworth M.F. (2009). The computation of social behavior. Science 324: 1160-1164.
13. Hampton A.N., Bossaerts P., and O'Doherty J.P. (2008). Neural correlates of mentalizing-related computations during strategic interactions in humans. Proc. Natl. Acad. Sci. U. S. A 105: 6741-6746.
14. Yoshida W., Dolan R.J., and Friston K.J. (2008). Game theory of mind. PLoS Comput. Biol. 4: e1000254.
15. Gold J.I. and Shadlen M.N. (2007). The neural basis of decision making. Annu. Rev. Neurosci. 30: 535-574.
16. Kable J.W. and Glimcher P.W. (2009). The neurobiology of decision: consensus and controversy. Neuron 63: 733-745.
17. Hare T.A., Schultz W., Camerer C.F., O'Doherty J.P., and Rangel A. (2011). Transformation of stimulus value signals into motor commands during simple choice. Proc. Natl. Acad. Sci. U. S. A.
18. Behrens T.E., Hunt L.T., Woolrich M.W., and Rushworth M.F. (2008). Associative learning of social value. Nature 456: 245-249.
19. Frith C. and Frith U. (2010). Learning from others: introduction to the special review series on social neuroscience. Neuron 65: 739-743.
20. Fleming S.M., Weil R.S., Nagy Z., Dolan R.J., and Rees G. (2010). Relating introspective accuracy to individual differences in brain structure. Science 329: 1541-1543.
21. Heekeren H.R., Marrett S., and Ungerleider L.G. (2008). The neural systems that mediate human perceptual decision making. Nat. Rev. Neurosci. 9: 467-479.
22. Chiu P.H., Kayali M.A., Kishida K.T., Tomlin D., Klinger L.G., Klinger M.R., and Montague P.R. (2008). Self responses along cingulate cortex reveal quantitative neural phenotype for high-functioning autism. Neuron 57: 463-473.
23. Kanai R., Bahrami B., and Rees G. (2010). Human parietal cortex structure predicts individual differences in perceptual rivalry. Curr. Biol. 20: 1626-1630.
24. Kanai R., Dong M.Y., Bahrami B., and Rees G. (2011). Distractibility in daily life is reflected in the structure and function of human parietal cortex. J Neurosci. 31: 6620-6626.
25. Kanai R., Bahrami B., Roylance R., and Rees G. (2011). Online social network size is reflected in human brain structure. Proc. Biol. Sci.
26. Campbell-Meikeljohn D., Kanai R., Bahrami B., Bach D.R., Dolan R.J., Roepstorff A., and Frith C.D. (2011). Structure of orbitofrontal cortex predicts social influence on value. Under review.
27. Gintis H., Bowles S., Boyd R.T., and Fehr E. (2005). Moral Sentiments and Material Interests: The Foundations of Cooperation in Economic Life. (Cambridge, MA: The MIT Press).
28. De Dreu C.K., Greer L.L., Handgraaf M.J., Shalvi S., Van Kleef G.A., Baas M., Ten Velden F.S., Van D.E., and Feith S.W. (2010). The neuropeptide oxytocin regulates parochial altruism in intergroup conflict among humans. Science 328: 1408-1411.