The Bayesian Infant and other Half-Baked Conceptualisations

Dans le cadre du séminaire de maîtrise de l'Orientation psychologie clinique, conférence de

Tobias Nolte, MD, MSc

Clinical Research Associate, Wellcome Trust Centre for Neuroimaging, UCL
Senior Researcher, The Anna Freud Centre
Visiting Lecturer, IPU, Berlin

Université de Genève, 18/12/2015

Disclaimer

- After this talk the listener may be completely confused with regards to the interchangeability of the concepts of attachment, mentalisation, epistemic trust and active inference

Overview

1. The Bayesian Infant and Epistemic Trust

2. Interpersonal Generative Models in Severe Psychopathology

3. The Computational Psychiatry Approach and some Preliminary Data

Acknowledgements

Peter Fonagy
Read Montague
Karl Friston
Michael Moutoussis
Christoph Mathys
Patrick Luyten
Andreas Hula
Yulien Huang
Jim Hopkins
Jennifer O'Connell

… and the entire PD study team and patients

Overview

1. The Bayesian Infant and Epistemic Trust

2. Interpersonal Generative Models in Severe Psychopathology

3. The Computational Psychiatry Approach and some Preliminary Data

Epistemic trust and secure attachment

- **Secure attachment** is created by a system that also induces a sense of **epistemic trust** ➔ that the information relayed by the caregiver may be trusted (i.e. learnt from)

- Evidence:
  - **Contingent** responsiveness to the infant's own (at first, automatic) expressive displays in secure attachment
  - During **"mirroring" interactions**, the other will "mark" her referential emotion displays in a 'manifestative' manner to instruct the infant
  - Cognitive **advantage** of secure attachment

(Fonagy & Allison, 2014, Corriveau et al., 2009)

Children's Trust in Mother's Claims

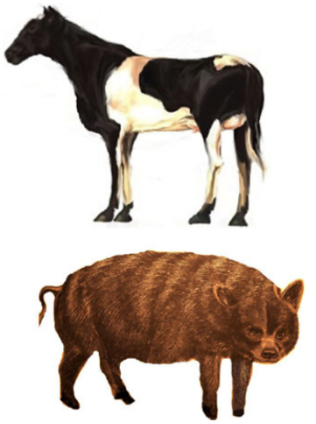


Figure 1. Examples of 50–50 hybrids (horse-cow and pig-bear).

Figure 2. Examples of 75–25 hybrids (bird–fish and rabbit–squirrel).

Proportion of Trials on Which Children Chose Their Mother by Task, Attachment Group, and Question

Question	Avoidant (n = 26)	Secure (n = 98)	Resistant (n = 9)	Disorganized (n = 16)
Condition 1: Novel objects				
Ask	.59 (.31)	.71 (.26)	.74 (.23)	.63 (.26)
Endorse	.57 (.28)	.63 (.27)	.75 (.29)	.62 (.24)
Condition 2: 50–50 pictures				
Ask	.52 (.24)	.66 (.21)	.64 (.18)	.56 (.21)
Endorse	.43 (.34)	.63 (.26)	.81 (.24)	.48 (.28)
Condition 3: 75–25 pictures				
Ask	.32 (.20)	.48 (.27)	.64 (.22)	.56 (.28)
Endorse	.48 (.27)	.40 (.27)	.53 (.29)	.48 (.25)
Total				
Novel objects	.58 (.26)	.67 (.24)	.74 (.20)	.63 (.20)
50–50 Pictures	.47 (.22)	.64 (.19)	.72 (.08)	.52 (.18)
75–25 pictures	.34 (.18)	.44 (.22)	.58 (.24)	.52 (.24)

Natural Pedagogy theory (Csibra & Gergely, 2006; 2009, in press)

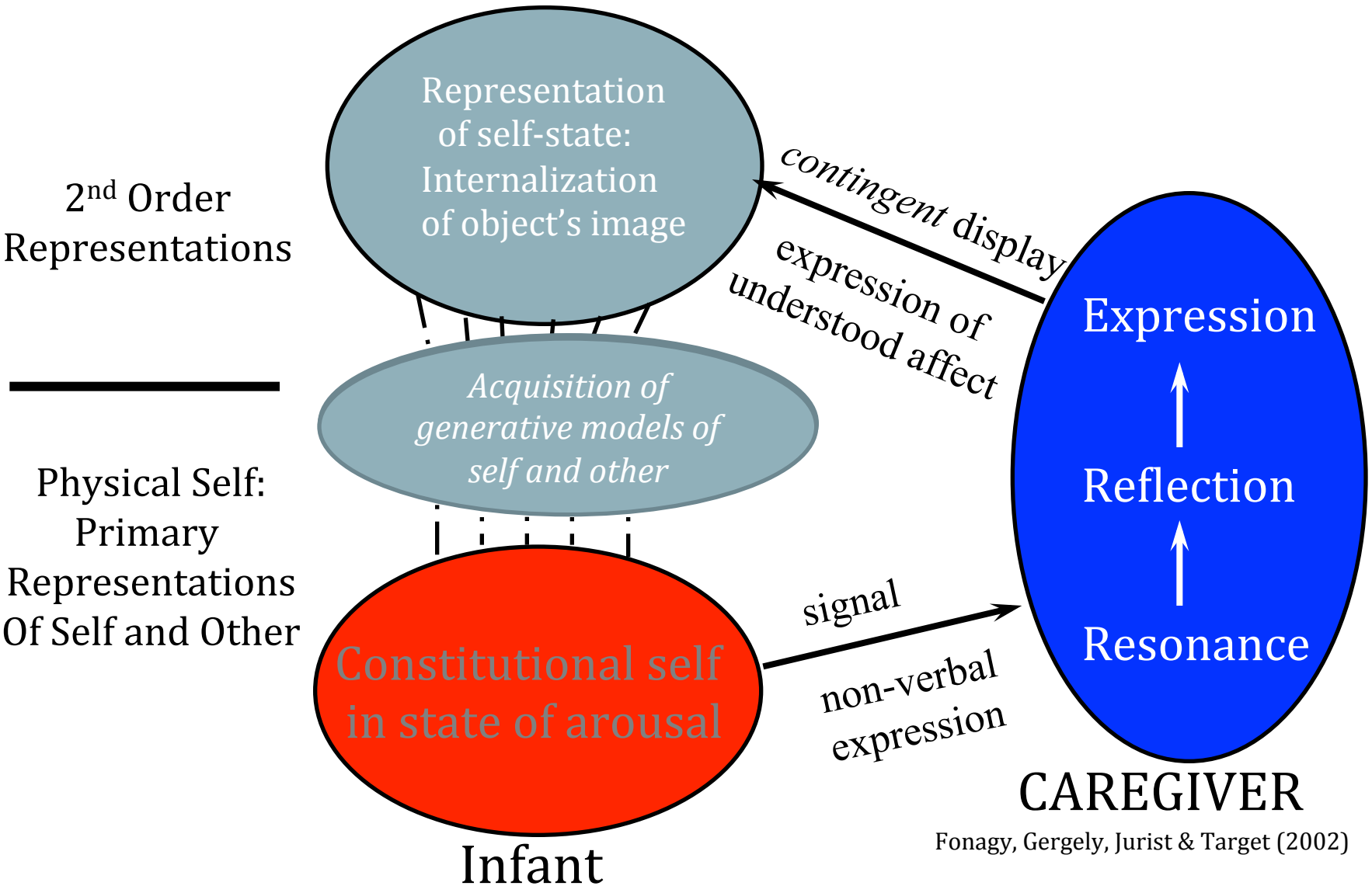
- A human-specific, cue-driven social cognitive adaptation of mutual design dedicated to ensure efficient **transfer of** relevant **cultural knowledge**
- Humans are predisposed to **'teach' and 'learn'** new and relevant cultural information from each other
- Human **communication** is specifically adapted to allow the transmission of
- a) cognitively **opaque** cultural knowledge
- b) kind-**generalizable** generic knowledge
- c) **shared** cultural knowledge

The Pedagogical Stance is triggered by Ostensive-Communicative cues

- Examples:
  - eye-contact, eye-brow flashing
  - turn-taking contingent reactivity
  - Motherese
  - being addressed by own name Csibra & Gergely, (2009, 2010)

- Ostensive cues function:
  - to signal that the other has a **Communicative Intention** addressed to the infant/child
  - to **Manifest New and Relevant** information about a referent

Transmission of Generative Models



Fonagy, Gergely, Jurist & Target (2002)

- The **mind is found within the other** not within itself
- Evolution has '**prepared**' our brains **for** psychological **therapy**
- We are **eager to learn** about the opaque mental world from those around us
- We are prepared to learn most readily about minds in **conditions of epistemic trust -> generalisability**
- **Therapy** is not just **about** the what but the **how of learning**
  - **Opening the person's mind** via establishing contingencies so (s)he once again can trust the social world by changing expectations

Mentalising

- Mentalizing is a form of ***imaginative*** mental activity **about** others or oneself, namely, perceiving, experiencing and interpreting human behaviour in terms of ***intentional*** mental states (e.g. needs, desires, feelings, beliefs, goals, purposes, and reasons).

- The capacity to mentalize has both “trait” and “state” aspects that vary in quality in relation to emotional arousal and interpersonal context, i.e. is dynamic within a set of constraints.

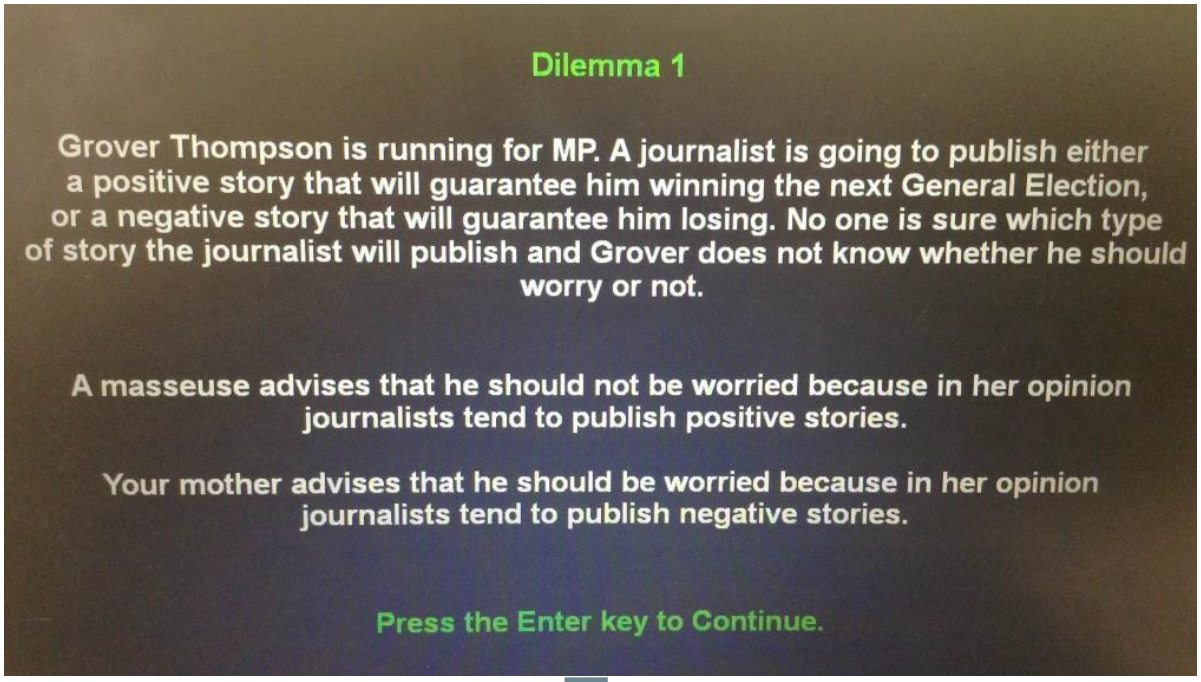
Data Excursion I – Epistemic Trust

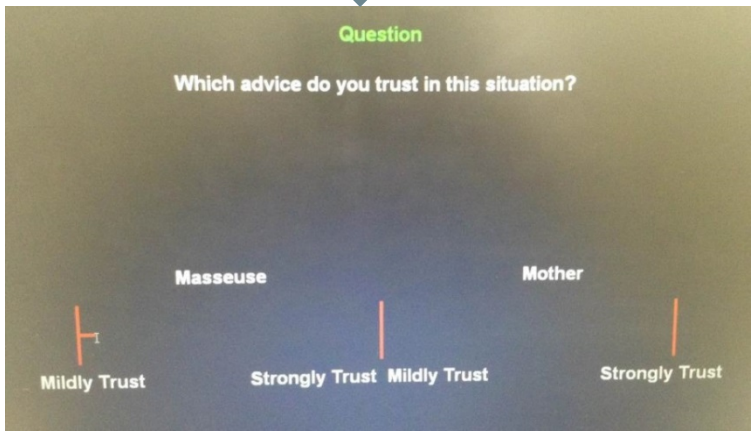
Aim of current ET study

... To develop a simple instrument that quantatively captures individual differences in how knowledge from others in contexts of uncertainty is trusted.

(O'Connell, Nolte & Fonagy, in prep.)

Measure


Dilemma 1
Grover Thompson is running for MP. A journalist is going to publish either a positive story that will guarantee him winning the next General Election, or a negative story that will guarantee him losing. No one is sure which type of story the journalist will publish and Grover does not know whether he should worry or not.
A masseuse advises that he should not be worried because in her opinion journalists tend to publish positive stories.
Your mother advises that he should be worried because in her opinion journalists tend to publish negative stories.
Press the Enter key to Continue.


Question
Which advice do you trust in this situation?
Masseuse
Mother
Mildly Trust
Strongly Trust
Mildly Trust
Strongly Trust

Measure

- 20 dilemma situations:

- Items adapted to be construed as impersonal dilemma.
- contain impersonal situations involving life-threatening topics, non-fatal health and safety risks, deceitfulness, job-loss, and stealing.
- ending with two opposing outcomes

(e.g., *There is a chance that those who take the vaccine will develop immunity to the deadly disease forever. Alternatively, there is a chance that those who take the vaccine will contract the disease instead*).

- typical “morally correct” response (e.g., not taking a wallet found on the ground) is also counterbalanced between sources

- Participants are asked to ignore their own opinions and assume that they are a blank slate with no clue about what is considered right and wrong by society.

Hypotheses

- HC participants will trust their mother's advice on significantly more occasions than the stranger's advice. On the contrary, it is conjointly predicted that patients with BPD will show no preference for either figure as a reliable source of information.

- Second, based on a continuous measure, it is predicted that HC individuals will have a significantly higher *strength* of epistemic trust for their mother than patients will do.

- Lastly, it it predicted that maternal attachment will be a significant predictor of epistemic trust for the mother and the stranger, but that maternal attachment will account for more of the variance in maternal trust than in stranger trust.

Sample

- One hundred and thirty nine individuals (41 men, 98 women, mean age 31.12, *SD* = 10.54, range = 17 – 55, IQ mean = 47.99, *SD* = 7.96, (18– 60) participated in the study. This included 96 patients with BPD and 43 healthy controls.

Table 1. *Demographic variables for conditions*

	Controls (n = 43)	Patients (n = 96)	Significance
Gender	23 males, 20 females	18 males, 78 females	$\chi 2$ (1, n = 139) = 17.23, p = .000
Age	29.86 (11.24)	31.69 (10.22) years	t (137) = .945, p = .347
IQ Score	49.72 (7.47)	47.16 (8.09)	t (132) = 1.75, p = .083

Results I - descriptives

Table 2. *Mean (SD) values for the main variables with respect to condition*

	Controls (n = 43)	Patients (n = 96)
Proportion Chose Mother	0.66 (0.20)	0.54 (.17)
Strength mother Trust	22.51 (9.20)	17.40 (8.17)
Strength stranger Trust	9.88 (7.55)	13.33 (7.59)
Confidence Average	33.60 (9.58)	34.54 (17.83)
ECR-R Avoidance	2.78 (1.17)	4.10 (1.43)
ECR-R Anxiety	2.90 (1.28)	5.10 (1.15)
IPPA	82.55 (21.86)	66.58 (23.31)
CTQ	41.51 (13.83)	65.50 (23.16)
PAI-BOR	24.13 (11.58)	54.37 (10.32)
GSI	0.53 (0.49)	2.08 (0.78)

Results II – bivariate correlations

Table 3. Bivariate correlations for main study variables

	2	3	4	5	6	7	8	9	10
1. Proportion chose mother	.727**	-.769**	.020	-.196	-.160	.508**	-.489**	-.255*	-.380**
2. Strength mother trust	-	-.333**	-.308*	-.164	-.073	.451**	-.388**	-.154	-.322**
3. Strength stranger trust		-	-.220*	.066	.182	-.395**	.424**	.256*	.269**
4. Confidence			-	.158	-.129	-.094	.042	.073	-.006
5. ECR-R Avoidance				-	.184	-.280**	.324**	.251*	.321**
6. ECR-R Anxiety					-	-.216*	.299**	.667**	.586**
7. IPPA						-	.657**	-.303**	-.460**
8. CTQ							-	.354**	.553**
9. PAI-BOR								-	.640**
10. GSI									-

** = p < 0.05, ** = p < 0.01*

Results III

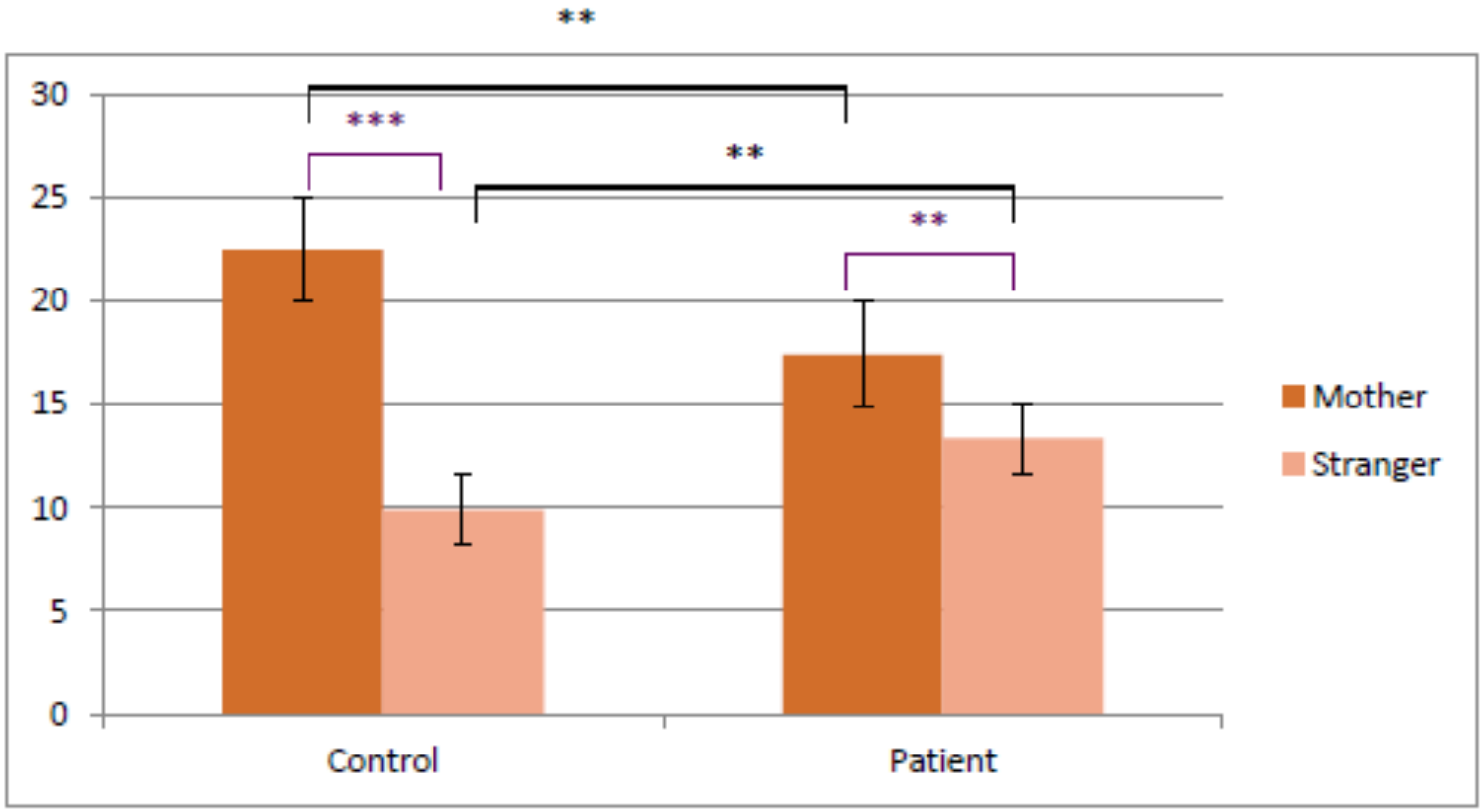


Figure 2. Comparison of trust strength between mother and stranger with respect to condition. * *p = 0.05*, ** *p = 0.01*, *** *p = 0.001*

Results IV

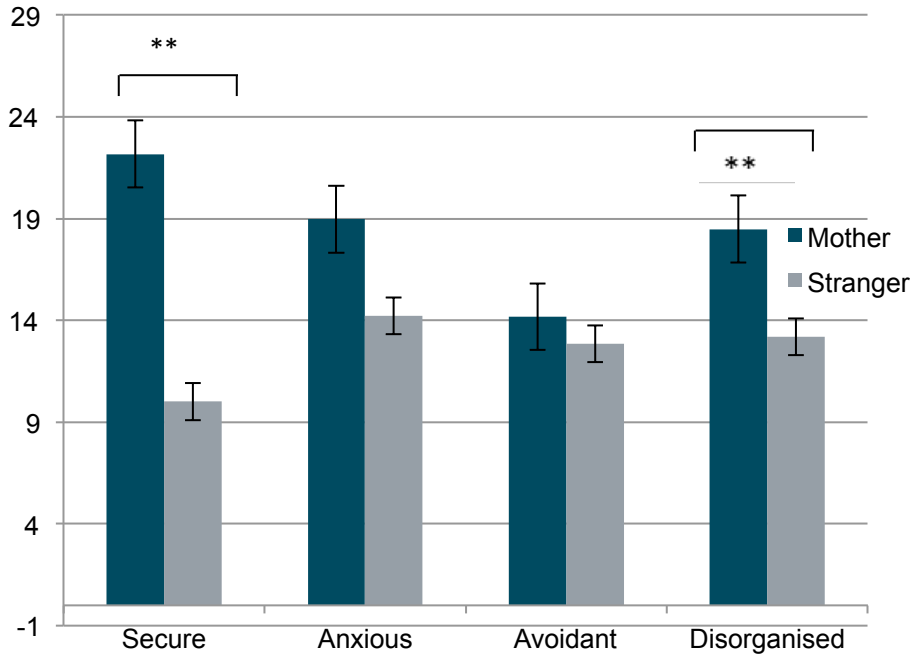


Figure 3. Comparison of trust strength between mother and stranger with respect to attachment style (ECR-R by median split) . ** p = 0.05, ** p = 0.01,*

Results V – hierarchical regression

Table 4. Summary of hierarchical multiple regressions

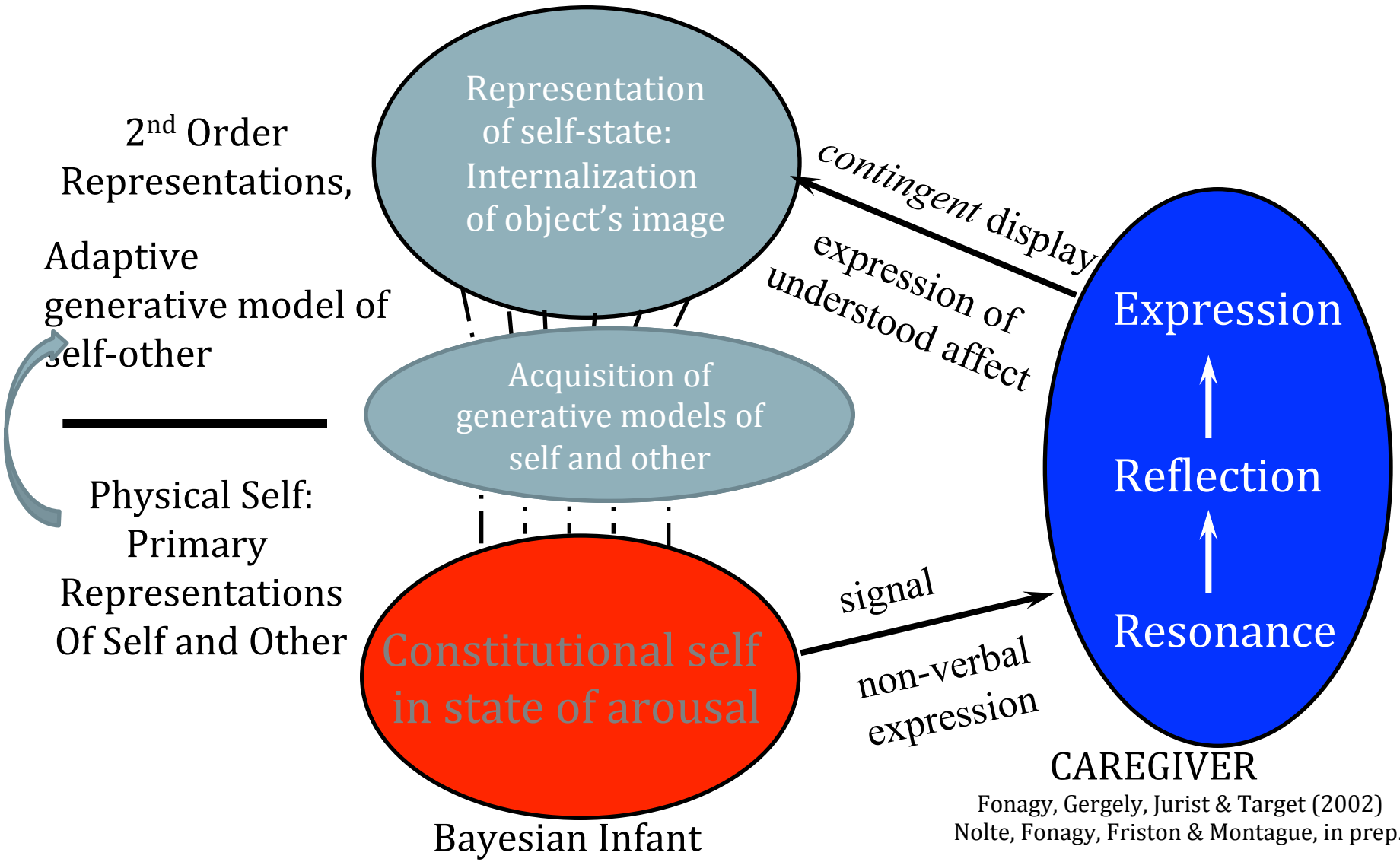
Outcome Variable	Step	Predictors	B	SE β	β	*t*
Maternal epistemic trust	1	Gender	-.805	1.83	-.04	-.44
		Condition	-4.83	1.81	-.26	-2.67
	2	Gender	.188	1.72	.01	.11
		Condition	-1.99	2.33	-.11	-.85
		Maternal attachment	.12	.04	.34	3.01**
		Childhood trauma	-.05	.05	-.12	-1.01
		GSI score	.12	1.14	.01	.11
Stranger epistemic trust	1	Gender	-1.42	1.63	-.08	-.87
		Condition	3.95	1.61	.24	2.46
	2	Gender	-2.10	1.53	-.13	-1.37
		Condition	.28	2.47	.02	.11
		Maternal attachment	-.05	.04	-.15	-1.28
		Childhood trauma	.12	.04	.36	2.83**
		Borderline traits	.04	.07	.09	.58
		GSI score	-.62	1.15	-.08	-.55

Note. * p < .05, ** p < .01

R squared change = .151,
F change (3, 135) = 7.36, p < .000

R squared change = .157,
F change (4, 135) = 5.56, p =< .000

Transmission of Generative Models



Fonagy, Gergely, Jurist & Target (2002)
Nolte, Fonagy, Friston & Montague, in prep.)

Bayesian Infant

Let’s imagine…

Basic principles underlying mentalisation/ Starting point for GMs of Self and Other

- Internal/mental states are **opaque**
- We only make **inferences** about them
- These are **prone to error**
- ET facilitates benign curiosity and *relational* training ground for learning
- Developmentally/Clinically:

  Overarching principle is to take an

  "inquisitive stance" **=**

  Interpersonal behaviour is characterised by an expectation that an individual's mind may be influenced, surprised, changed or even enlightened by learning about another mind

Bayesian Infant

„If indeed the brain is a Generative Model of the world then much of it must be occupied by modeling other people.“

„I would go as far as saying up to 95% of its computations.“

Freud

"A hungry baby screams or kicks helplessly. But the situation remains unaltered, for the excitation arising from an internal need is not due to a force producing a momentary impact but to one which is in continuous operation. A change can only come about if in some way or other (in the case of the baby, through outside help) an 'experience of satisfaction' can be achieved which puts an end to the internal stimulus.

An essential component of this experience of satisfaction is a particular perception (that of nourishment, in our example) the mnemic image of which remains associated thenceforward with the memory trace of the excitation produced by the need. As a result of the link that has thus been established, next time this need arises a psychical impulse will at once emerge which will seek to re-cathect the mnemic image of the perception and to re-evoke the perception itself, that is to say, to re-establish the situation of the original satisfaction."

Freud, Standard Ed., Vol. V, p. 565

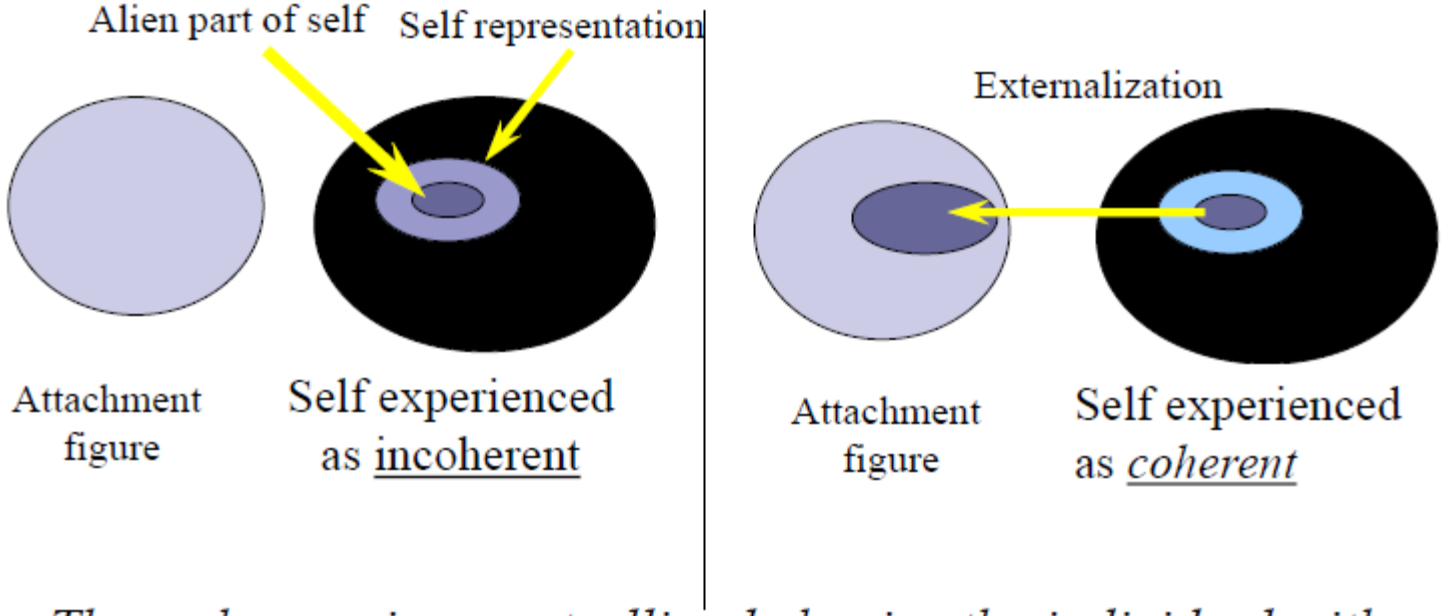
Overview

1. The Bayesian Infant and Epistemic Trust

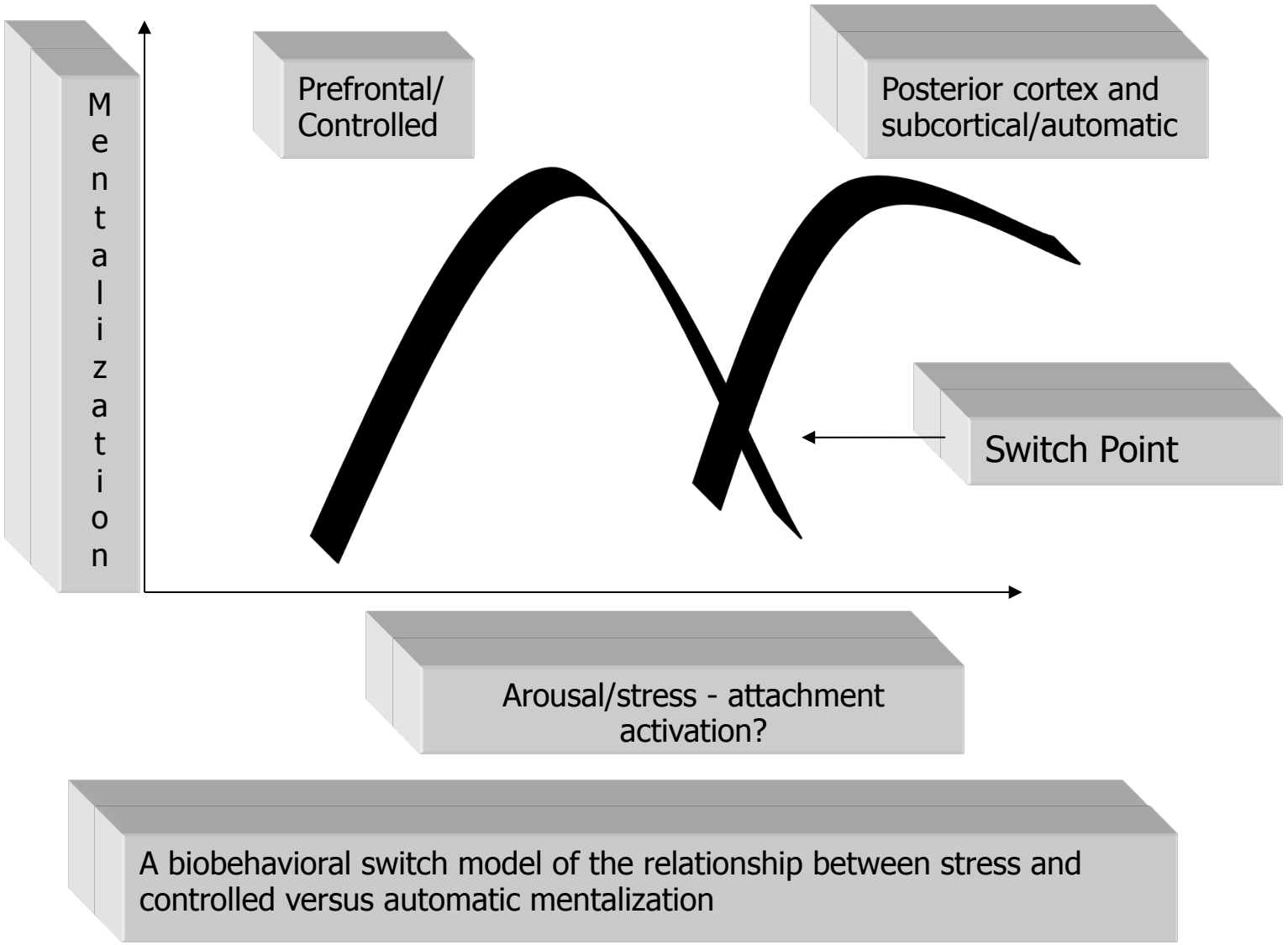
2. Interpersonal Generative Models in Severe Psychopathology

3. The Computational Psychiatry Approach and some Preliminary Data

Attempts to create a coherent sense of self via “projective identification” under hightened attachment arousal



Through coercive, controlling behavior the individual with disorganized attachment history achieves a measure of coherence within the self representation



(Fonagy & Luyten, 2009;
Nolte et al., 2013)

Effects of Relational Trauma – Epistemic Distrust

Disorganised attachment experiences result in:

Hallmark: pronounced interpersonal prediction errors

- A pervasive pattern of instability of interpersonal relationships, self-image, and affect, and marked impulsivity etc.
- Key features: impaired and markedly dysfunctional interpersonal functioning underpinned by characteristic deficits in mentalising self and/or others
- Common (etiological) pathway: inflexible, sub-optimal generative model of interpersonal dynamics rooted in early attachment experiences
- May account for inter-individual variability and the non-normative and non-optimal deviations from Bayesian computations
- At the neural level: reinforcement of synaptic gains that encode relational expectations

(Fonagy & Luyten, 2009;
Nolte, Fonagy, Montague & Friston, in prep.)

Fonagy:

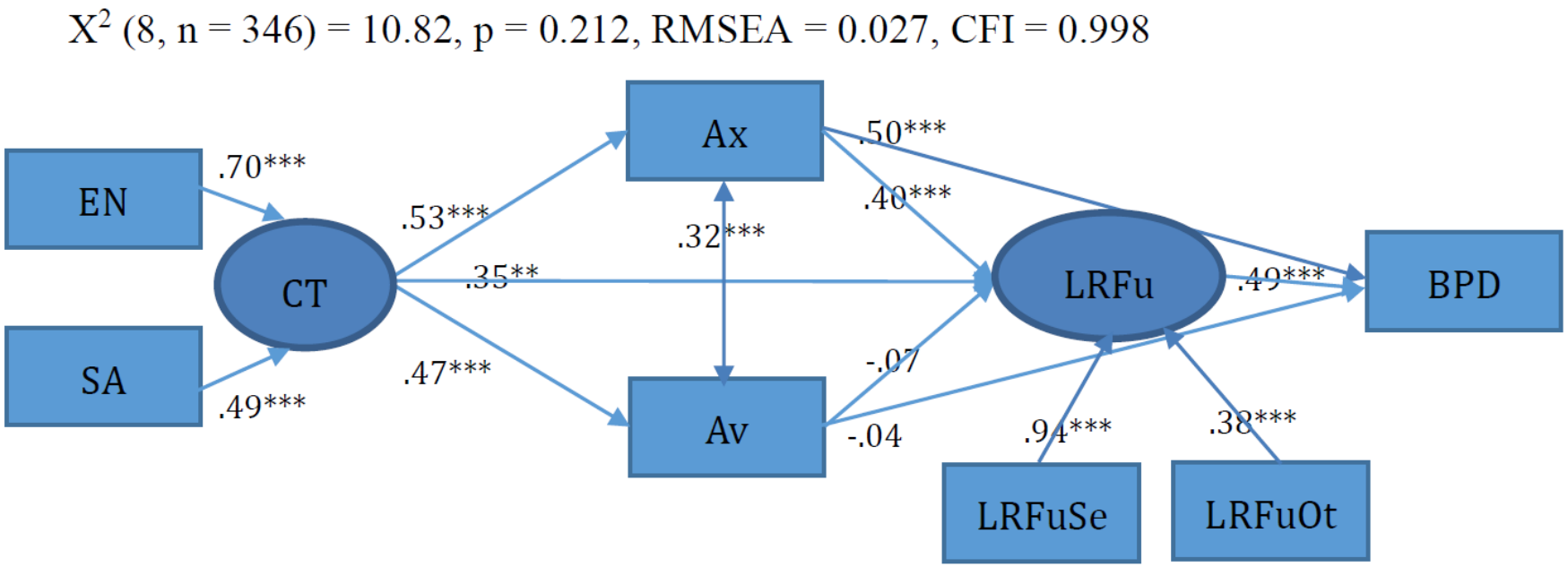
"We have to see *the destruction of trust in social knowledge as the key mechanism in pathological personality development.*

BPD, therefore, is not a "personality" disorder; it involves a state of social inaccessibility. It can be conceived of as a temporary state of incompatibility with an evolutionarily designed intracultural communication system. It describes a state of isolation from communication from one's partner, one's therapist, one's teacher—all resulting from epistemic mistrust and hypervigilance."


I hear you
but I'm not
listening

Data Excursion II – SEM for vulnerability

SEM for pathways from adversity to vulnerability for PD

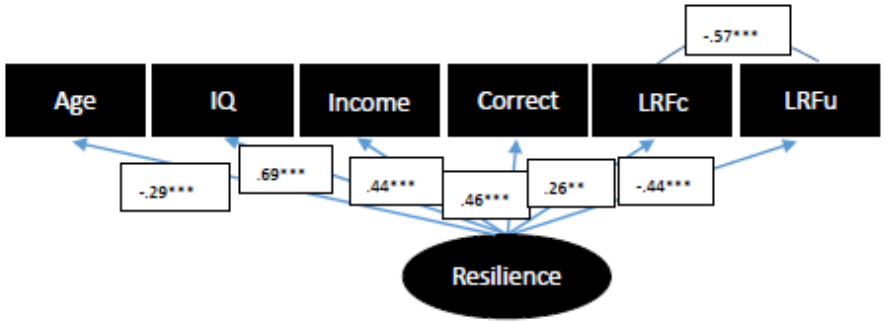


(Nolte, Huang, Zimmermann & Fonagy, in prep.)

Resilience

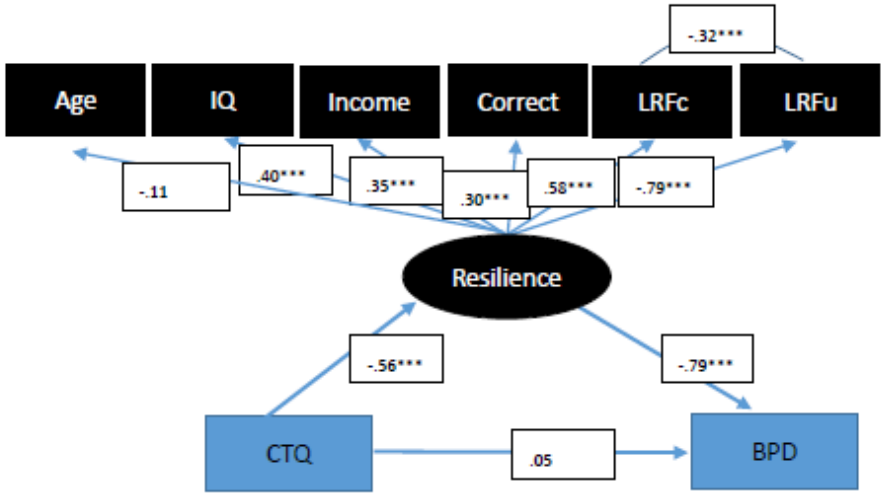
8. CFA

X^2 (8, n = 298) = 7.43, p = .491, RMSEA = 0.000, CFI = 1.000



9. SEM

X^2 (18, n = 298) = 81.06, p < .000, RMSEA = 0.109, CFI = 0.885



Overview

1. The Bayesian Infant and Epistemic Trust

2. Interpersonal Generative Models in Severe Psychopathology

3. The Computational Psychiatry Approach and some Preliminary Data

Limitations of current cognitive endophenotype approaches and DSM-5

General (content-independent) cognition

- Nonspecific cognitive processes clearly play a role but are not sufficient explanations

  - Much of **cognitive processing** is inherently content (**semantics**) **dependent** and this could affect even supposedly ‘‘neutral’’ information processing ‘‘tasks’’

  - **Not all** patients cognitively **compromised** (on classical measures)

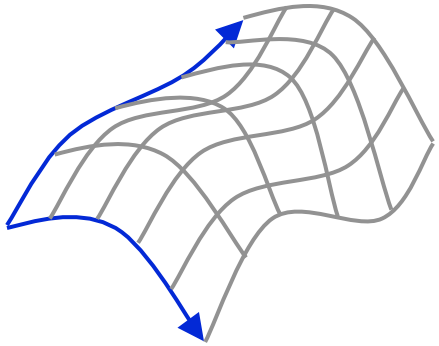
  - Significant **proportion** of cognitively **compromised** subjects do **not** experience **symptoms**

Computational phenotyping: How can computation connect mind variables to brain variables?

Growing body of cellular & molecular data

Computational
Models and methods

Behavior, thoughts, moods, etc.



Describe complex behaviors, thoughts, moods, etc as **computations**.

↕

Describe neuronal interactions as **computations**.

Our approach

1. Use a collection of economic probes to extract a vector of parameters (and brain responses) to characterize healthy populations.

2. Compare these with clinical samples of Borderline Personality and Anti-Social Personality Disorder to identify neurobiological commonalities and distinct neural patterns, also: pre and post treatment.

3. Attempt to detect phenotypical markers that cut beyond polythetic categorical DSM approach.

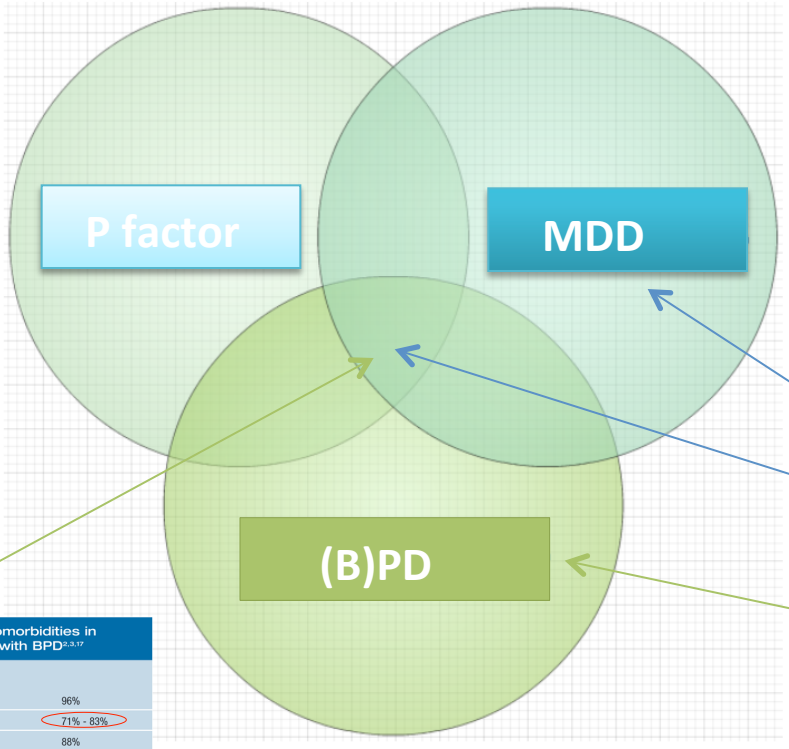
4. Demonstrate clinical utility in predicting symptom change, drop out and to improve triage pathways.

Lots of underwriting results (UCL Social Exchange Battery, PI: R. Montague)

The social exchange battery	Modeling results on which to 'lean'
1. Ultimatum game with 'norm changes'	1. Bayesian observer-defined norms
2. Multi-round trust game (10 rounds)	2.1 Depth-of-thought 'typing'/Investors 2.2 Level of Trust and Coaxing /Trustee
3. Bargaining game (buyer)	3. Strategic style 'typing'
4. Bargaining game (seller)	4. Model of interpersonal suspiciousness
5. Social Hierarchy Task	5. Dominance/Aggression
6. Observed Trust Game (70 rounds)	6. Model for mentalising of another "relationship"

Established corresponding neural signatures: Xiang, Lohrenz & Montague, 2013,
King-Casas, Lohrenz, Fonagy & Montague, 2008, Bhatt et al., 2010; 2012, King-Casas et al., (in prep.)
For conceptual overview: Montague, Dolan, Friston & Dayan, 2012)

Proposed Extension



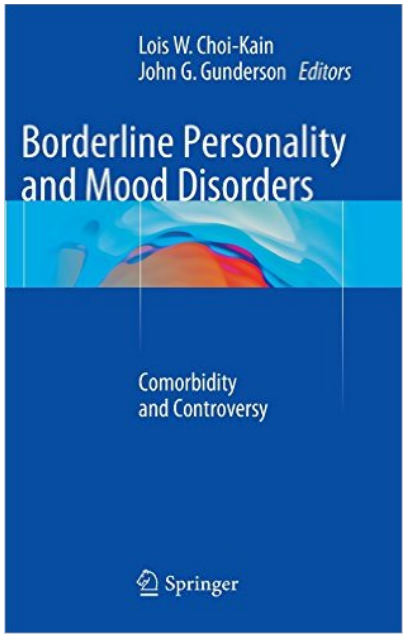


Table — Lifetime comorbidities in patients with BPD[2,3,17]

Most common Axis I comorbidities	
Any mood disorder	96%
MDD	71% - 83%
Any anxiety disorder	88%
Panic disorder	34% - 48%
PTSD	47% - 56%
Alcohol or substance use disorder	50% - 65%
Any eating disorder	7% - 26%
Most common Axis II comorbidities	
Avoidant personality disorder	43% - 47.4%
Obsessive-compulsive personality disorder	18.2% - 25.7%
Dependent personality disorder	16% - 50.7%
Paranoid personality disorder	13.7% - 30.3%

Completed:
N=404
of which 199 scanned

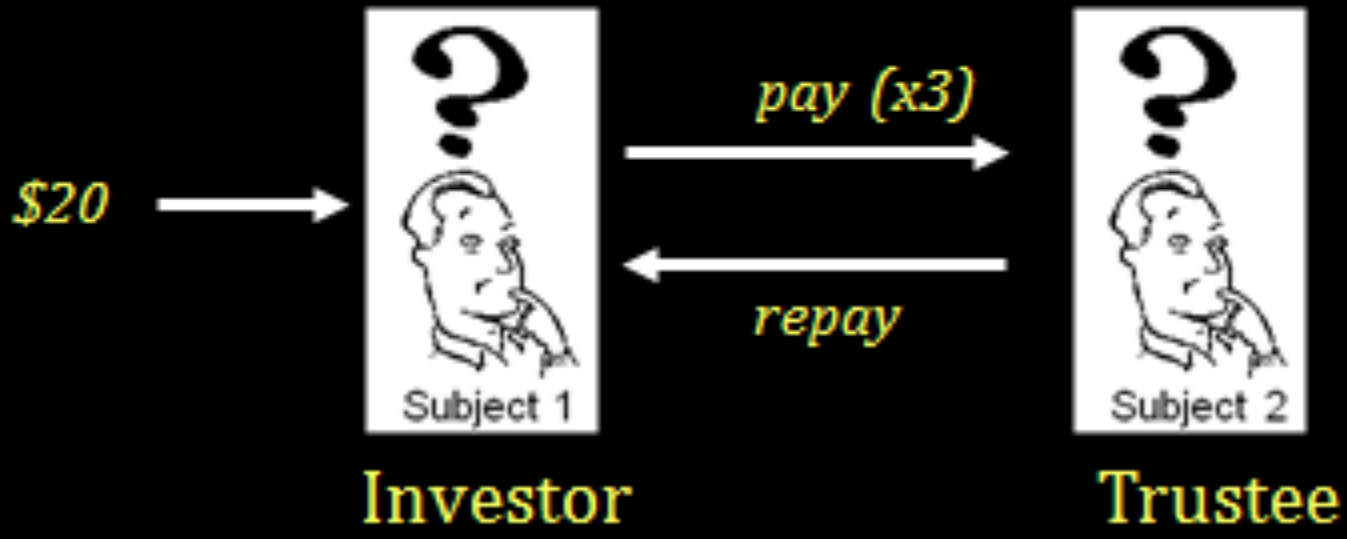
The multi-task approach generally asks two questions

1. How do variables relate across tasks and within a subject?

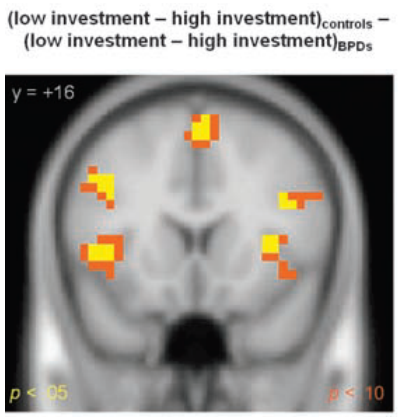
2. How do the vectors vary across clinical populations and in contrast to HCs, especially in areas of expected overlap?

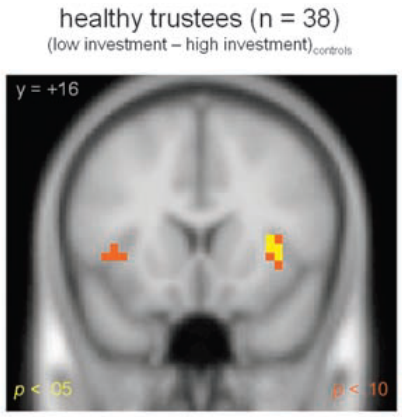
Probing reciprocity and model-building with a 10-round 'trust' game

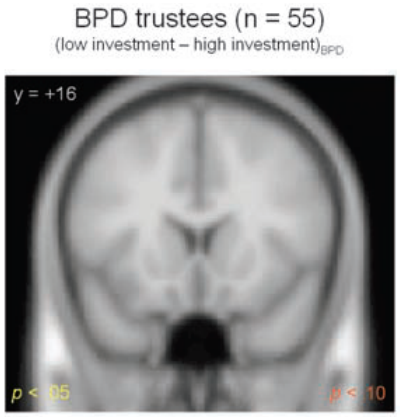
(Guth et al., 1982 → Weigelt and Camerer, 1988 → Dikhaut et al., 1995)



Evidence for differential patterns of BPD TG

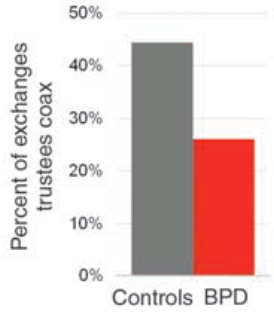


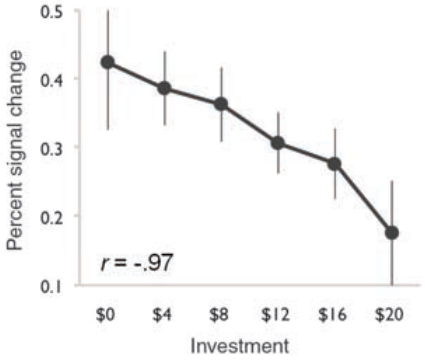


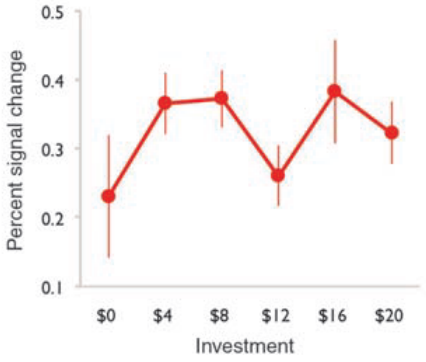


Region-of-Interest Analysis:

- percent BOLD signal change
- bilateral anterior insula (115 voxels identified by group difference)
- healthy and BPD trustees matched on sex, age, verbal IQ, and education



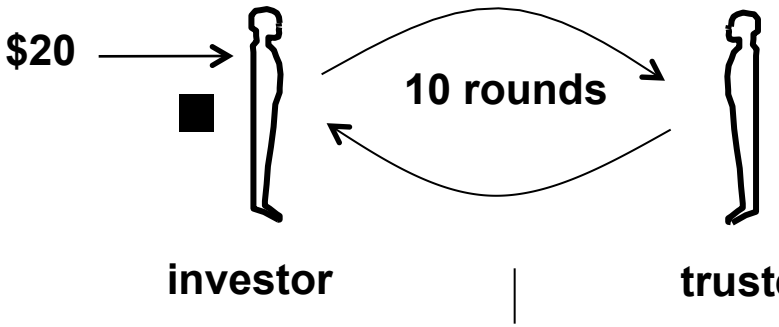




King-Casas et al., 2008

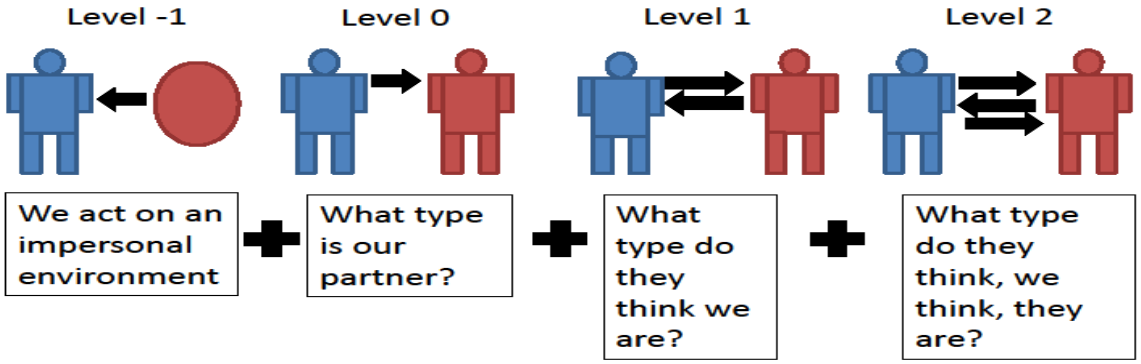
Theory-of-mind based **classification**

multi-round trust game



n = 178 pairs, trustee played by bot

Computational Theory of Mind



Ray D, *et al.* Bayesian model of behaviour in economic games. *NIPS* (2008)
Hula, Dayan & Montague (2015)
Nolte*, Hula*, Dayan, Fonagy & Montague (in preparation)

How do BPD and HC differ in the TG on a number of computationally derived parameter estimates?

- Theory of Mind
- Horizon Planning over Time

Data Excursion III – recursive modelling in two-part interactions

executing action a at state s (sometimes writing r_t for the reward following the t^{th} action). The expected reward given history h is

$$\mathcal{R}_h^a = \mathcal{R}(a, h) := \mathbb{E}[r|a, h]. \tag{3.5}$$

The discount function weights the present impact of a future return. We use a slightly unusual form, allowing it to be conditional on the present and future histories. The full discounted expectation of the reward r_τ at time τ (with history h_τ) from the perspective of the present time t (with history h_t, a prefix of h_τ) is thus

$$\mathcal{R}(a, h_t, h_\tau) := \gamma(h_t, h_\tau)\mathbb{E}[r_\tau|a_\tau = a, h_\tau]. \tag{3.6}$$

However, here, we use simpler discounting, depending only on the separation between present and future

$$\gamma(h_t, h_\tau) = \gamma(\tau - t), \qquad h_t \subset h_\tau, \tag{3.7}$$

and is strictly 0 if $h_t \not\subset h_\tau$. If all $\gamma(\tau - t)$ are 0 for $\tau - t > P$, we say that the local planning horizon at h_t is less or equal than P.

The policy $\pi \in \Pi$

$$\pi(a, h) := \mathbb{P}[a|h] \tag{3.8}$$

is defined as a mapping of histories to probabilities over possible actions. Here Π is called the set of admissible policies. For convenience, we sometimes write the distribution function as $\pi(h_t)$.

The value function of a fixed policy π starting from present history h_t is

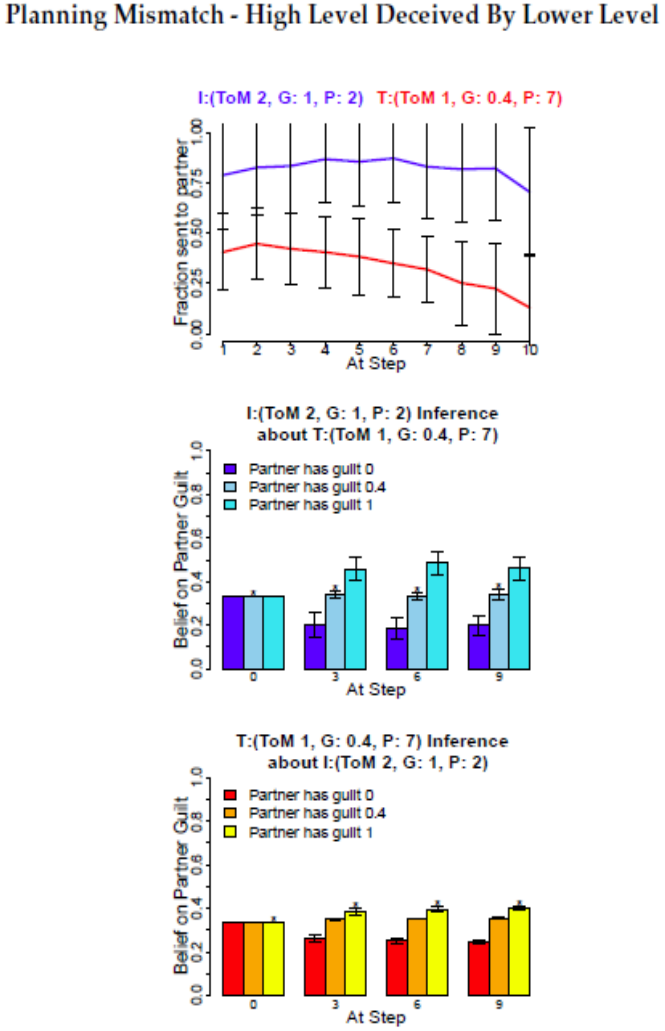
$$V^\pi(h_t) := \sum_{\tau=t}^{\infty} \gamma(\tau - t)\mathbb{E}[r_\tau|a_\tau \sim \tilde{\pi}(h_\tau)] \tag{3.9}$$

as a sum of the discounted future expected rewards (note that h_τ is a random variable here, not a fixed value). Equally, the state-action value is

$$Q^\pi(a, h_t) := \mathcal{R}(a, h_t) + \mathbb{E}[\sum_{\tau=t+1}^{\infty} \gamma(\tau - t)\mathbb{E}[r_\tau|a_\tau \sim \pi(h_\tau)]. \tag{3.10}$$

Our results so far

- Hula, Montague & Dayan (2015)
  - POMCP to estimate horizon planning
  - Current work on irritability/ (volatility)



Examples from simulated data based upon the model

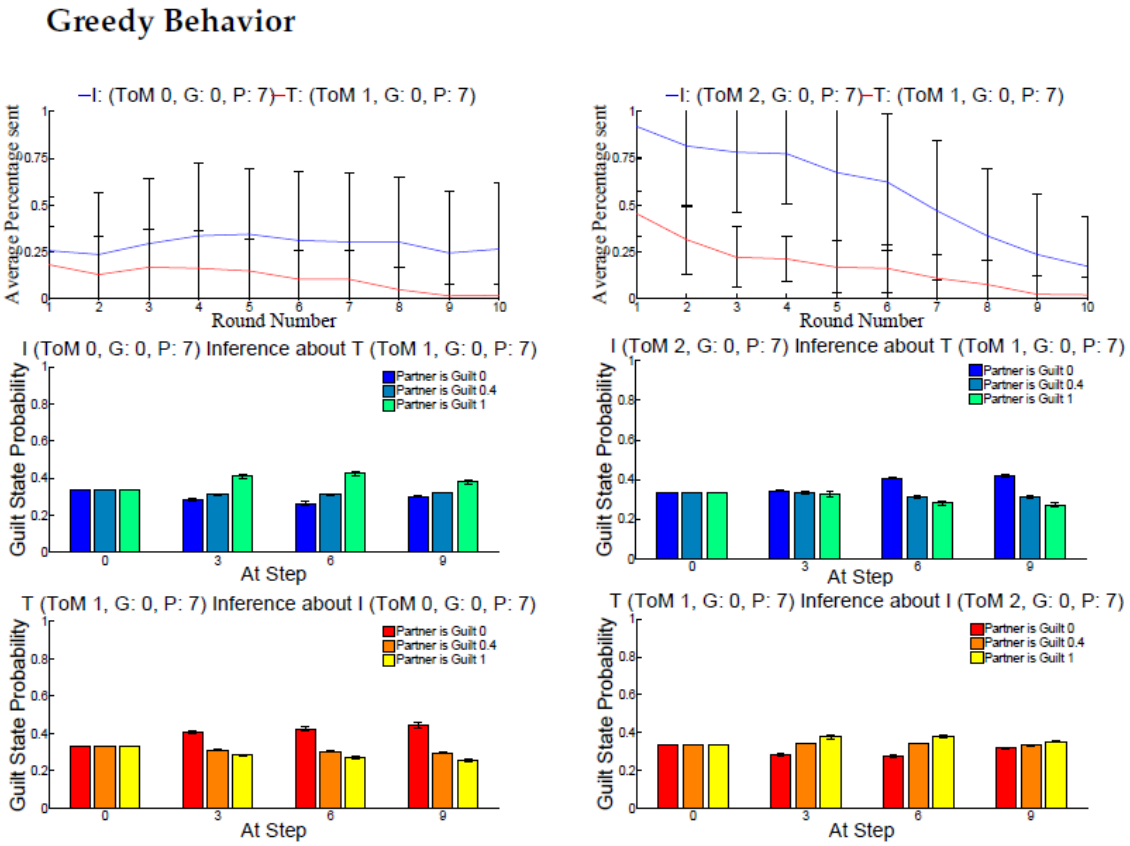


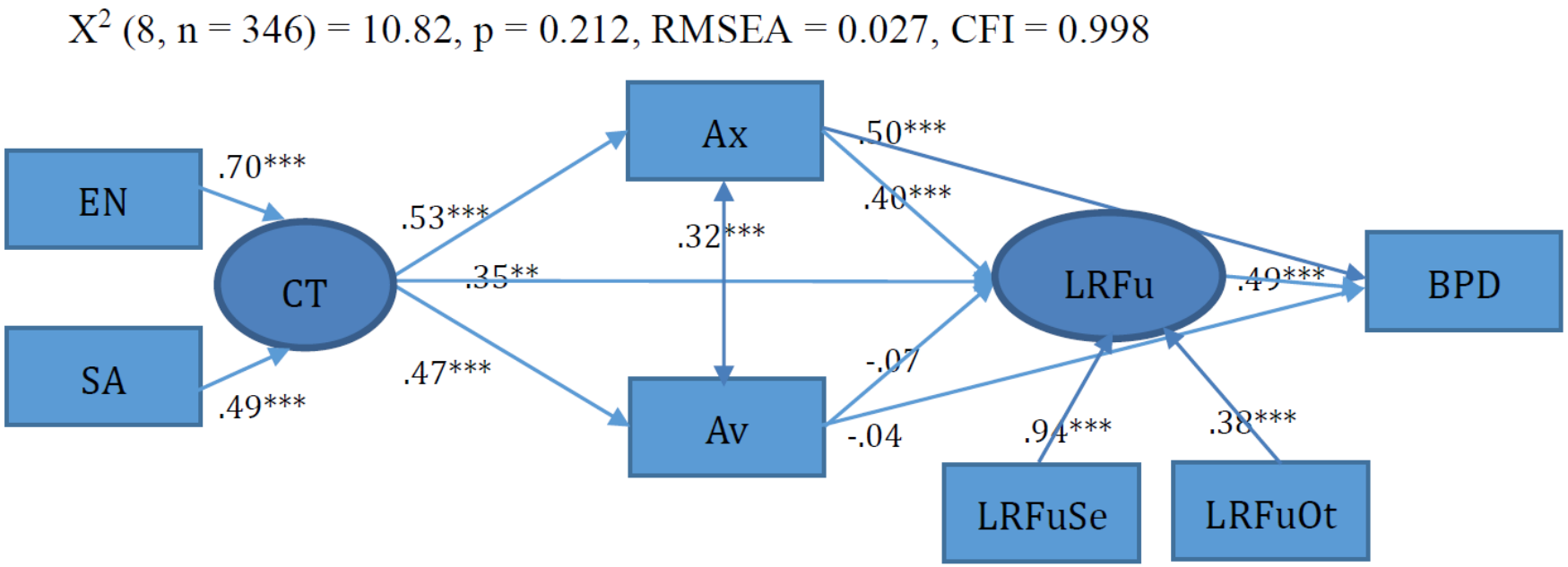
Figure 8: Averaged Exchanges (upper) and posteriors (lower). Left plots: Investor $(k^I, \alpha^I, P^I) = (0, 0, 7)$; Trustee $(1, 0, 7)$; right plots: Investor $(2, 0, 7)$ and Trustee $(1, 0, 7)$. The posterior distributions are shown for $\alpha = (0, 0.4, 1)$ at four stages in the game.

Our results so far

- Nolte*, Hula* et al. (in prep)

- Disease states probed as effect of role in iterative exchanges/recursive modelling in TG

SEM for pathways from adversity to vulnerability for PD



(Nolte, Huang, Zimmermann & Fonagy, in prep.)

Summary

Development of generative models: Organising principle for evaluation of early experiences

Level of epistemic trust ensures adaptation to one's developmental niche -> Resilience

Infant's experience of self and others is organised according to Bayesian updating of priors

Computational models of interpersonal functioning allow for mapping of developmentally compromised phenomena onto brain processes (mechanistic understanding)

These are characterised by model optimisation to reduce surprise/free energy – failure due to developmental adversity requires complexity reduction to explain away prediction errors,

GM are too narrow to predict cooperation/new/benevolent experiences

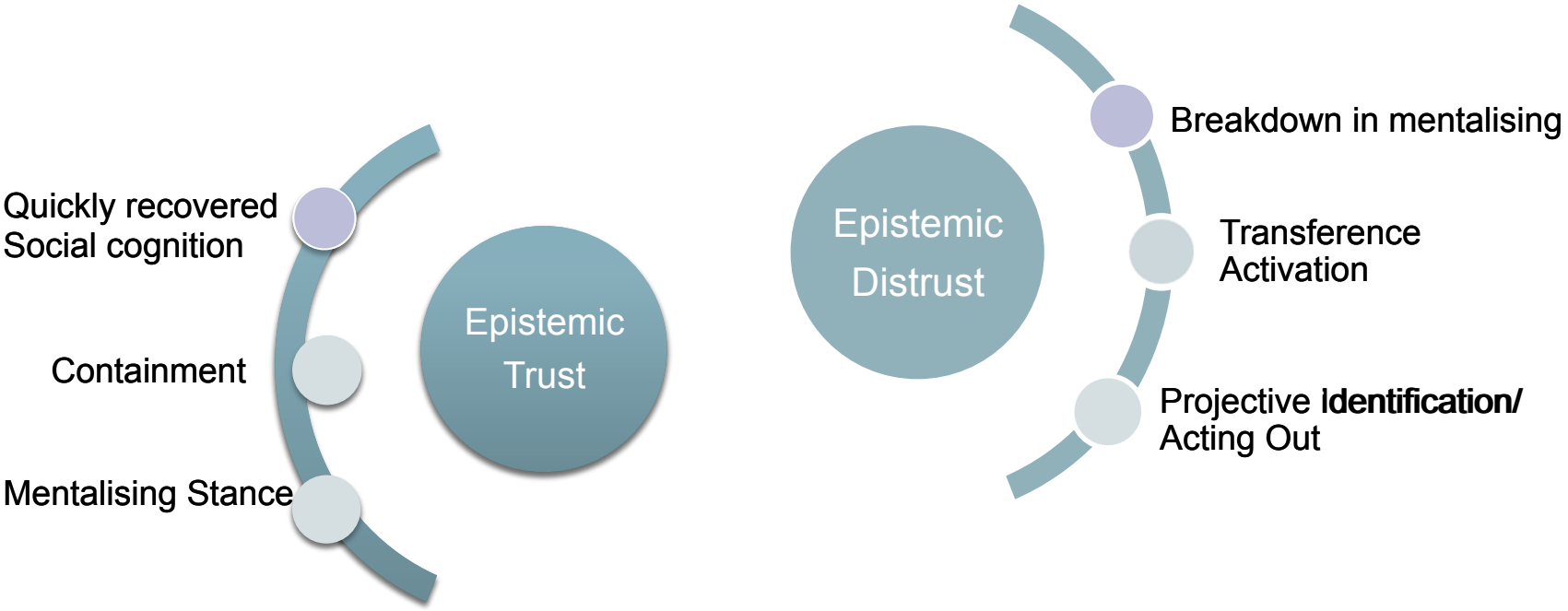
Prospective longitudinal needed to track GMs over time and in relation to normal/ pathological development

Clinical Implications

- Epistemic trust and the therapeutic relationship
- Rupture/repair processes
- Titrating of interventions
- Hard to reach patients

Attachment system activation

(e.g. mourning a loss, a devastating diagnosis, challenge during an assessment...)



Thanks!

t.nolte@ucl.ac.uk