

# Technisierte *moral agents*?

Mögliche Wechselwirkungen zwischen einer  
„Neurobiologie der Moral“ und dem  
therapeutischen Einsatz von  
Neurotechnologien

Markus Christen,  
Graduiertenprogramm UFSP Ethik,  
Universität Zürich

# Übersicht

## **Einführung:**

Einbettung des Projektes  
Fragestellung

## **Neuroscience of Ethics (NoE):**

Grundproblem und Motive  
*Was ist neuroscience of ethics / das wissenschaftliche Umfeld*  
Moral simplified / Moralische Stimuli / Korrelate moralischer Kognition

## **Tiefenhirn-Stimulation (DBS):**

Anwendungsfelder  
Assessment  
Berührungspunkte DBS-NoE

## **Moral agency:**

*Agency / moral agency / Ebenen von agency*  
Beeinträchtigung von *agency*  
*Agency* als Thema der Erfolgskontrolle von DBS?

## **Problemfelder:**

*Moral agency* testen  
*Moral agency* verändern

# Einbettung (1)

UFSP-Ethik Graduiertenprogramm Universität Zürich: 15 (+1) Projekte, davon 5 (+1) Projekte im weiteren Umfeld von Neurowissenschaften, Verhaltensforschung und Ethik.

**An evolutionary approach of morality: Social norms in a captive chimpanzee group**

**Goals:** The existence and enforcement of social norms are considered to be uniquely human features. This PhD-project picks up this assumption and investigates whether our closest living relatives, the chimpanzees, possess evolutionary precursors of morality (e.g. moral emotions and empathy) and thus may be capable to form and even enforce social norms in their groups.

**Methods:** A set of experiments will test for the existence of social norms in a captive chimpanzee group. These experiments include showing the animals short videos of possible norm violations and the monitoring of their reactions towards them. Behavioural observations will accompany and complete these experiments.

**Contact:** Claudia Rudolf von Rohr: rudo@rohr@ethik.uzh.ch  
**Affiliation:** Anthropological Institute, Prof. Canal van Schaik  
**Funding:** UFSP Ethik, Universität Zürich

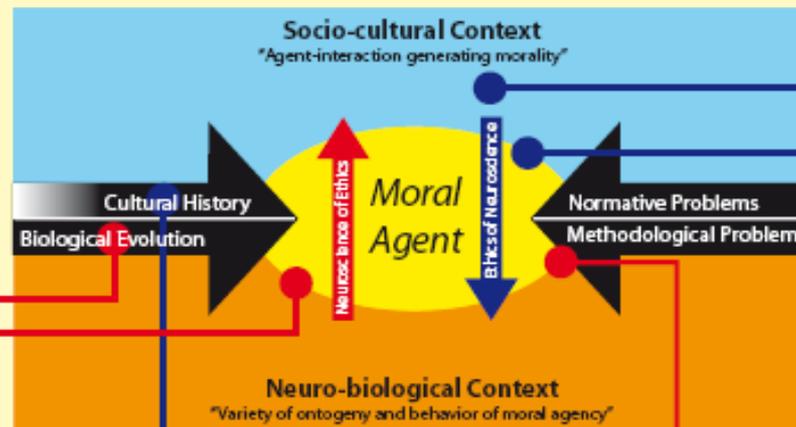


**The Biology of Moral Behavior: Motivational and evolutionary bases of altruism and cooperation**

**Goals:** The aims of this project are to investigate why humans behave altruistically and cooperative on a motivational and evolutionary level. Furthermore, since cooperation and altruism have been suggested as the very basis of human morality, the implications for a naturalistic moral theory are discussed.

**Methods:** Food sharing and cooperation experiments will be conducted with captive chimpanzees in a way that allows direct comparison with economic games played with humans. Similarities in the mechanisms underlying cooperation and altruism among humans and chimpanzees would suggest a biological foundation of human morality.

**Contact:** Adrian Jaeggi: jaeggi@ethik.uzh.ch  
**Affiliation:** Anthropological Institute, Prof. Canal van Schaik  
**Funding:** Cognition Foundation, Zürich

**The neurobiology of moral agency: Some specifications from the perspective of philosophy and cultural science**

**Cultural history of the "moral brain"**

**Goals:** Investigating the history of neuroscience in the 20th century along three narratives – the holistic brain (early 20th century with focus on Constantin von Monakow), cybematic brain (after WW II), and the enhanced brain (late 20th, early 21st century). Focusing on continuities and discontinuities in notions about brain functions and correlation with dominant ideas in public discourse.

**Methods:** Foucault's concept of governmentality and technologies of the self; Fleck's theory of thought styles.

**Contact:** Svenja Matswil: svenjamatswil@idm.darm.ac.uk

**Philosophy of "neuralized" moral agency**

**Goals:** Investigating the possibilities and problems of a neuro-scientific investigation of moral agency. Reconstruction of the current empirical models of moral agents within neuropsychology and comparison to philosophical concepts of moral agency. Evaluating possible neuroethical consequences of the "neuroscience of ethics" in the context of the identification of "moral pathologies".

**Methods:** Analysis of clinical records on moral pathologies related to frontal lesions. Focus on methodological problems of neuralized moral agency.

**Contact:** Marius Christen: christen@ethik.uzh.ch

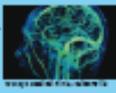
**Affiliation:** Zurich University Centre for Ethics, Prof. Johannes Fischer; Chair for Science Studies, ETH Zürich, Prof. Michael Wagner; Neuropsychology Unit, University Hospital Zurich, Prof. Marianne Regard  
**Funding:** Swiss National Science Foundation

**The application of neuroscience in economics: A concept of Neuroethics for Neuroeconomics**

**Goals:** This PhD-project is focusing on the intersection of economics and neuroscience. Within this new discipline called neuroeconomics, brain imaging technologies and knowledge about brain mechanisms are used to advance economics in theory and practice. We are interested in a critical reflection of neuroeconomics and concern about emerging ethical risks. The aim is to meet these risks with an appropriate ethical concept.

**Methods:** Normative concepts of economic theory, approaches of business ethics and neuroethics; evaluation of neuroeconomic research projects and practical examples.

**Contact:** Ina Kaufmann: kaufmann@ethik.uzh.ch  
**Affiliation:** ISU, Prof. Andreas G. Scherer  
**Funding:** UFSP Ethik, Universität Zürich



**Ethical problems raised by neuroenhancement: A systematic review and validation**

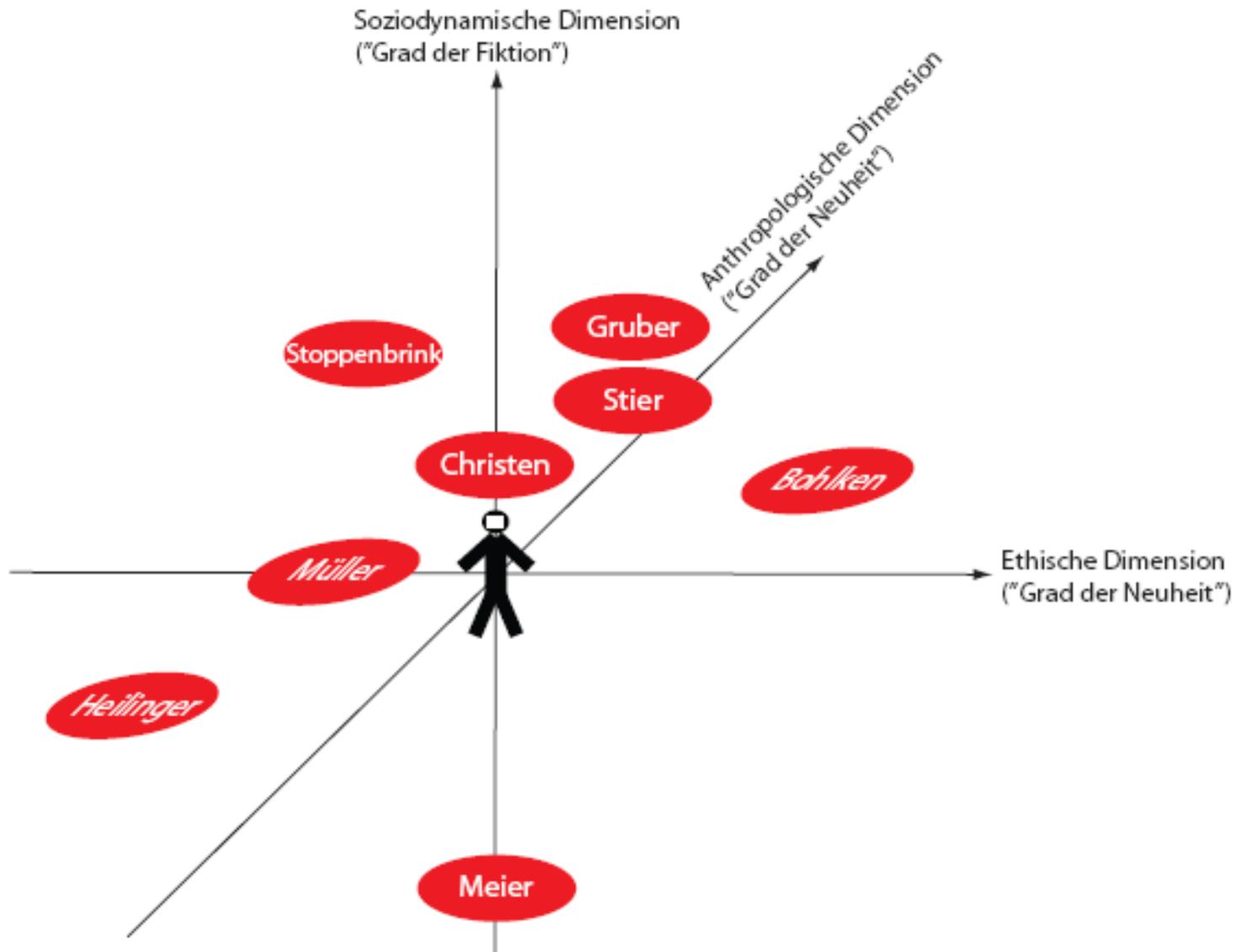
**Goals:** The variety of neuroenhancing interventions and their practicability in everyday clinical practice are constantly growing. Research on wider options is far from having reached its limits. The goal of the project is to identify the ethical issues raised by "neuro-enhancement" with special regard to pharmacologically based interventions and to discuss the positions held in the debate, focusing on their underlying normative assumptions.

**Methods:** Literature-review, interviews with experts and potential "users" of neuroenhancing interventions.

**Contact:** Nicole Miller: miller@ethik.uzh.ch  
**Affiliation:** Institut für Biomed. Ethik, Prof. Nikola Biller-Andero  
**Funding:** UFSP Ethik, Universität Zürich



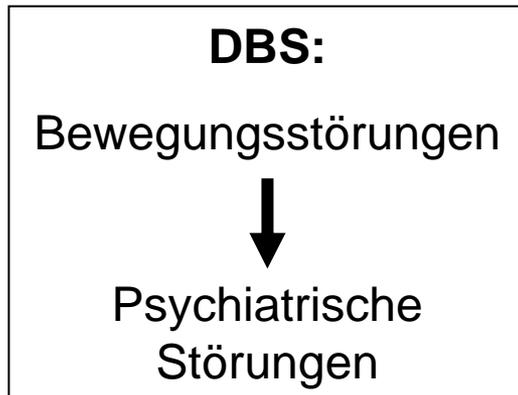
# Einbettung (2)



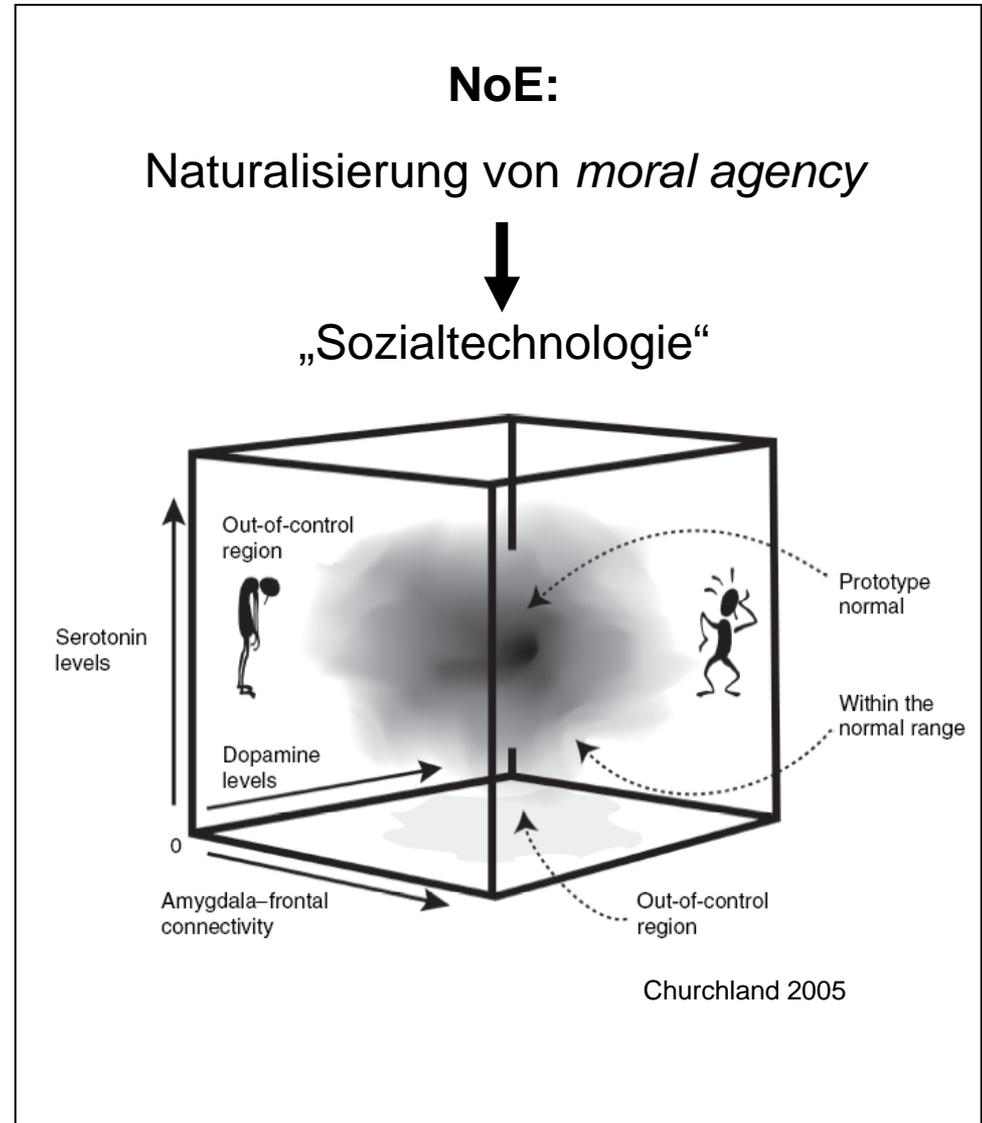
# Fragestellung (1)

Immediately after stimulation her attitude changed remarkably. From the first week, she was less anxious and felt more relaxed and easy-going. Her parents conclude: 'If she had made a pilgrimage, we would have thought a miracle had happened'. Some remaining compulsions hardly interfered with social activities and she was able to control and resist them a lot better. Relief of symptoms was markedly clear and reproducible by switching DBS on. With stimulation she was very active during the day. She reported feeling more cheerful and bustling, experiencing more excitement, talking louder and faster and laughing more. She sometimes seemed disinhibited, but did not engage in risky, reckless or impulsive behaviours.

# Fragestellung (2)



- Inhaltliche Konvergenz?
- Vergleichbare Kernprobleme?

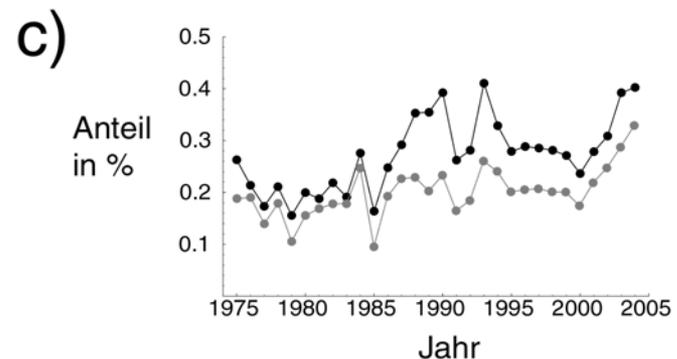
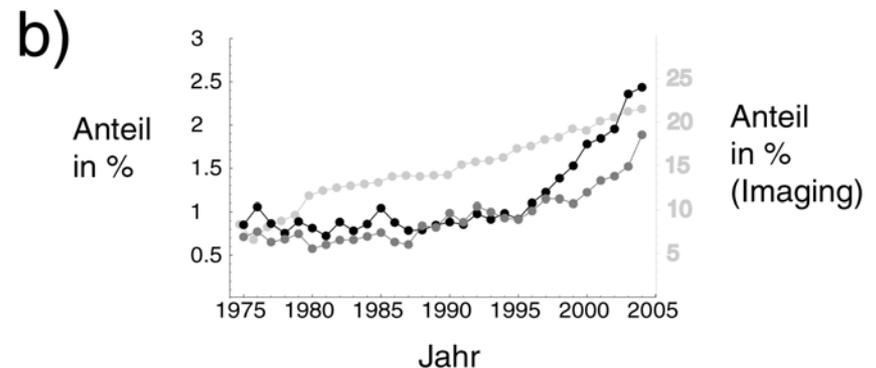
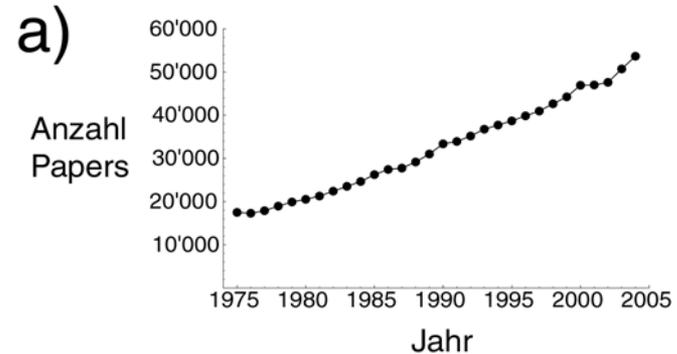


# Die *Neuroscience of Ethics* (NoE)

## Etwas Bibliometrie:

Basis: MedLine (1975-2005)

- Genereller Anstieg der Publikationstätigkeit in der Neurowissenschaft.
- Relativer Anteil der Bereiche *Imaging*, *Emotion* und *social cognitive neuroscience*.
- Relativer Anteil der Arbeiten über ethische und moralische Aspekte in der Neurowissenschaft (mit und ohne Transplantation von Hirngewebe).



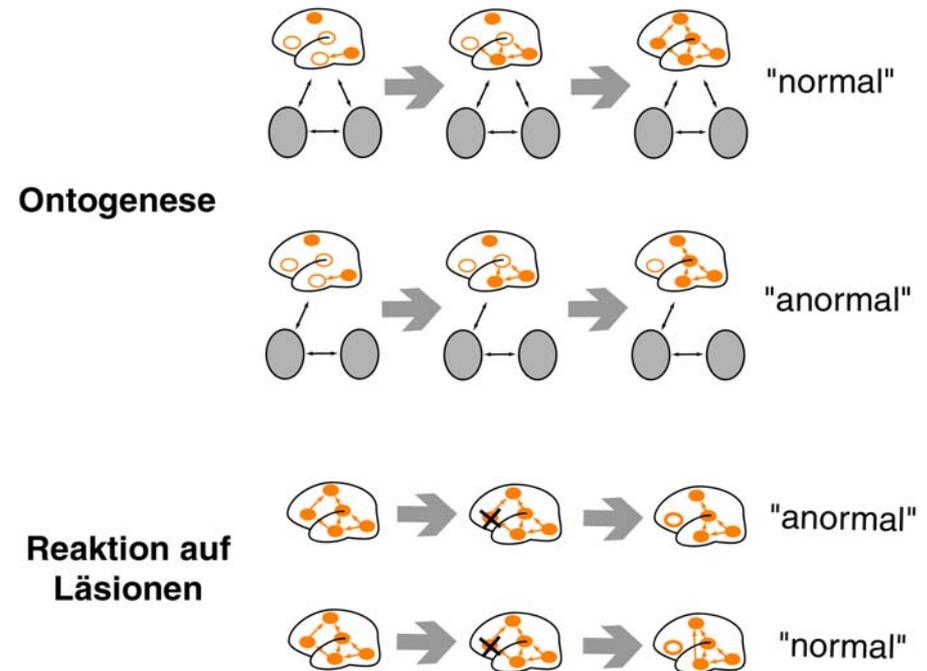
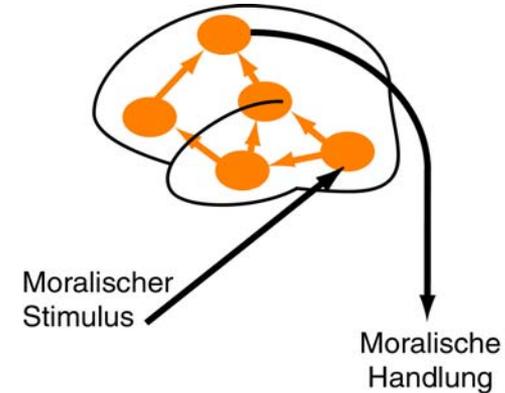
# Das naturwissenschaftliche Umfeld

- Social cognitive neuroscience*: Neuronale Grundlagen des Sozialverhaltens, *decision making*, *social pain*.
- Spiegelneuronen: Potentielles neuronales Substrat für Imitationsverhalten, Empathie.
- Experimentelle Ökonomie: Analyse von Handlungsstrategien, Motiven
- Emotionsforschung: Rolle von Emotionen beim Verhalten, *moral emotions*
- Primatologie: Moralähnliche Verhaltensweisen, Vorformen von Moral
- Neurologie/Neuropsychologie: „Moralische Pathologien“ und deren neuronale Basis.

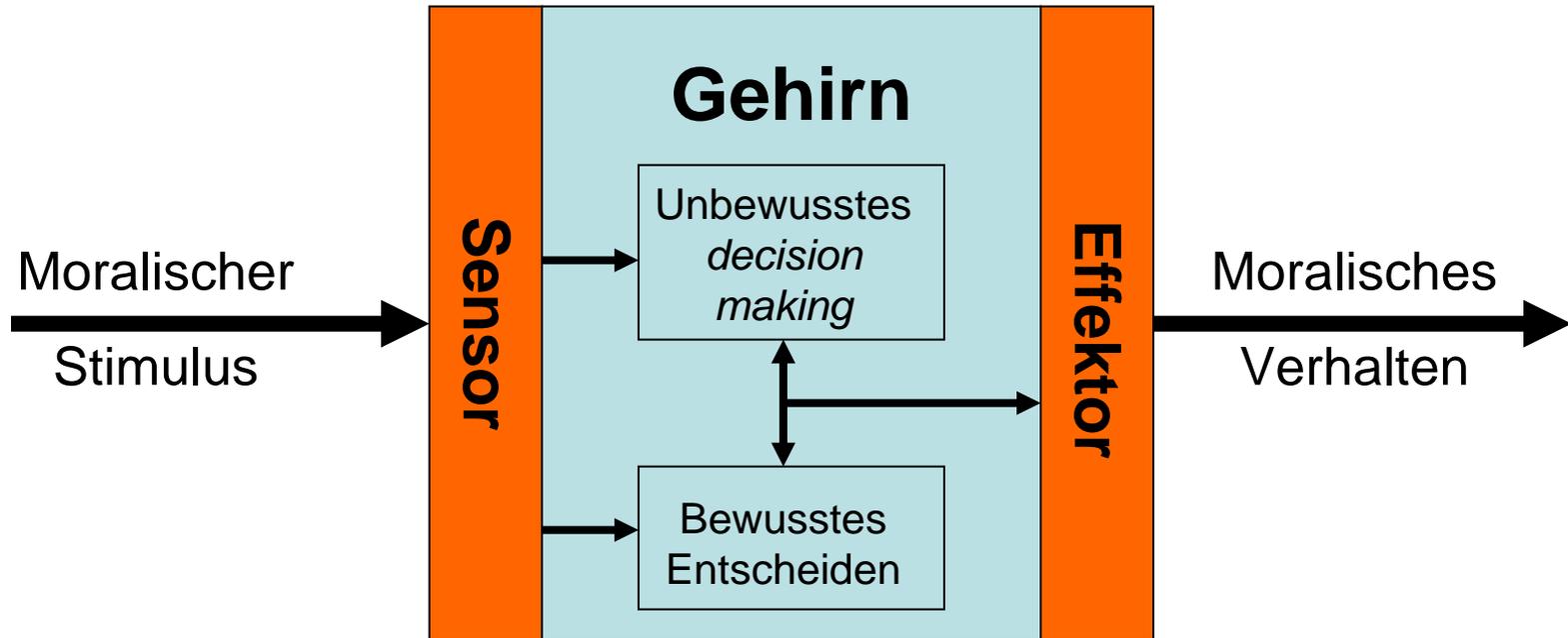
# Zum Ziel einer NoE

(Ein) Ziel der Erforschung der neurobiologischen Grundlagen von *moral behavior*.

- Ermittlung der funktionellen Netzwerks der „Moralfähigkeit“
- Untersuchung der Ontogenese dieses Netzwerkes (innere und äussere Einflüsse)
- Untersuchung der Reaktion des neuronalen Netzwerkes auf Schädigungen („moralische Plastizität“)



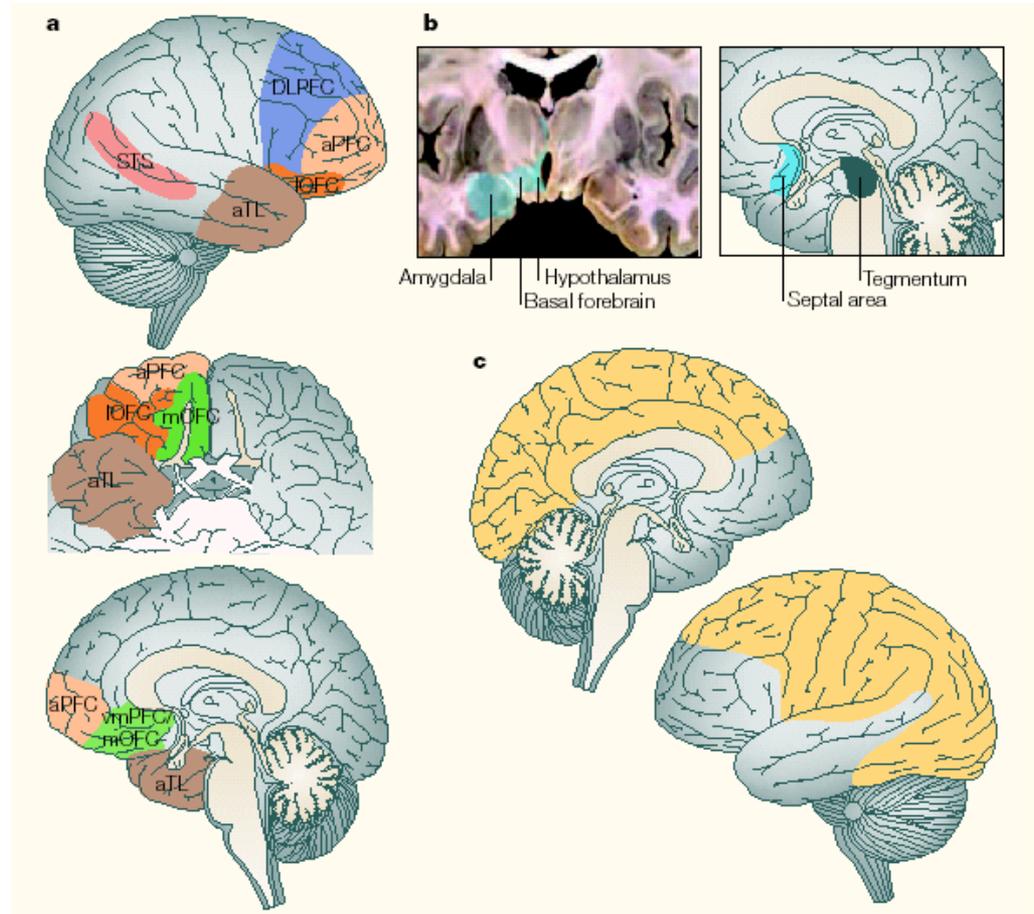
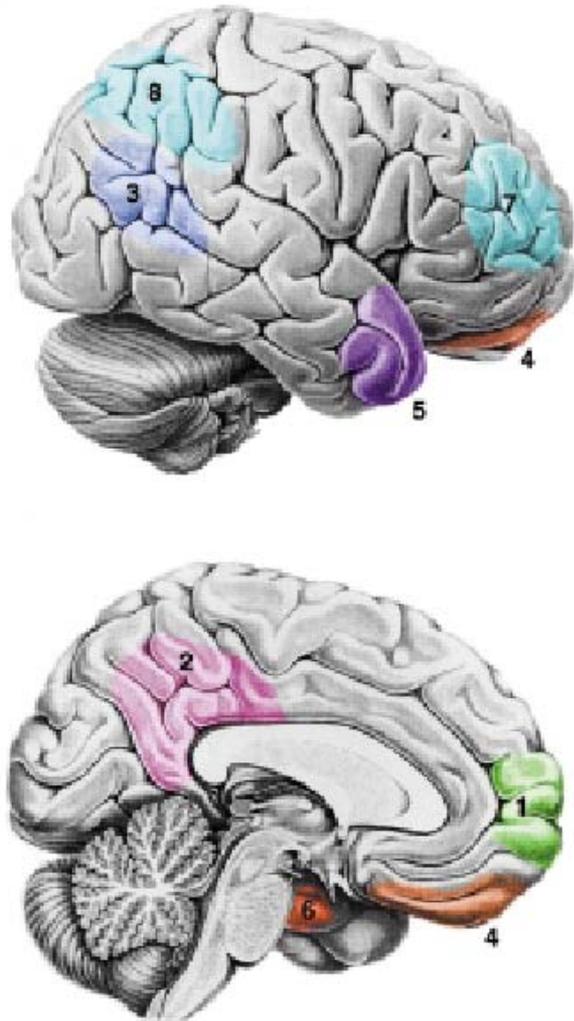
# Ein einfacher Blick auf *moral behavior*



# Moralische Stimuli (vorab fMRI)

- Bilder:** Meist Fotografien von Menschen „portraying emotionally charged, unpleasant social scenes, representing moral violations“. Problem: unerkannte Korrelationen zwischen Bildern (z.B. Gewaltdarstellungen)
- Sätze:** Meist mit moralisch einfach zu bewertenden Sachverhalten. Problem: Sind „*Baseline*-Sätze“ frei von moralischer Konnotation?
- Dilemmas:** *Personal vs. Impersonal* Dilemmas (Greene) (d.h. gewissermassen Grad des körperlichen Einbezugs). Problem: Was genau misst man, da der Entscheidungsprozess sich über viele Sekunden erstreckt.

# Neuronale Korrelate moralischer Kognition



# Tiefenhirnstimulation: Anwendungen

- Historische Wurzeln gehen sicher zurück in die Mitte des 20. Jahrhundert (Penfield, Delgado).
- In der heutigen Form: Seit den 1980er Jahren (Therapie von Bewegungsstörungen).
- Derzeit rund 35'000 Personen mit DBS (bis Ende 2006, Hardesty & Sackeim 2007).
- In den letzten Jahren: Experimentelle Anwendungen für weitere Störungen:
  - Angststörungen
  - Autismus
  - Depression
  - Tourette-Syndrom
  - Zwangsstörungen
  - ...

# DBS: Funktionsweise und Erfolgskontrolle

- Der genaue **Mechanismus** der DBS auf zellulärer Ebene ist weiterhin unklar („Modulation“ neuronaler Netze).
- **Effekte:** z.T. sofort eintretend (z.T. transient), z.T. erst nach einem gewissen Zeitraum einsetzend.
- **Stimulationsort-Bestimmung:** Erfahrungen der Psychochirurgie, Modelle, *trial and error* (McIntyre et al 2004)
- Statistisch gesicherte Erfahrungen hinsichtlich **Erfolg** vorab im Bereich Bewegungsstörungen:
  - In der Mehrzahl der Fälle (klare) Besserung der Symptome.
  - (Eher) geringe Beeinflussung anderer Aspekte (z.B. Kognition)
  - Z.T. starke Einzelfall-Abweichungen.
  - Kontroverse um Suizidrisiko (2004-2006).
  - Kritik an Qualität von „Quality of life“-Untersuchungen

# DBS und NoE: Berührungspunkte

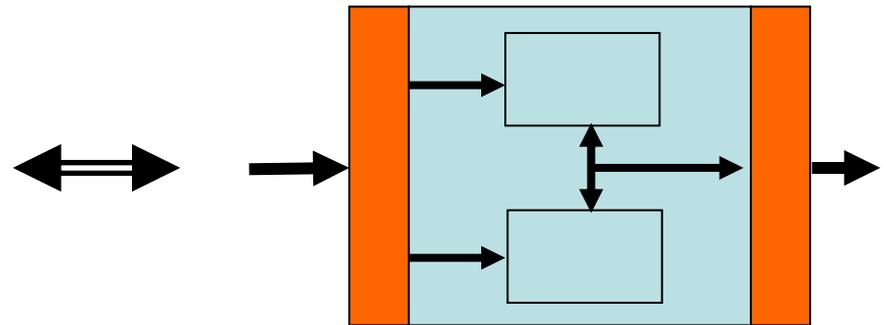
- **Stimulationsort:** Offenbar wird DBS an Orten im Gehirn angewendet, welche die NoE (in unspezifischem Sinne) als „wichtig“ für *moral cognition* identifiziert.
- **Evaluation:** Beide Ansätze sind mit dem Problem konfrontiert, ein komplexes Phänomen (*moral agency*, bzw. *quality of life*) halbwegs quantitativ fassbar zu machen.
- **Agency:** Letztlich brauchen beide Ansätze eine genauere Bestimmung von „Agency“, zumal dies ein Kernaspekt der oben genannten Phänomene ist.

# Agency

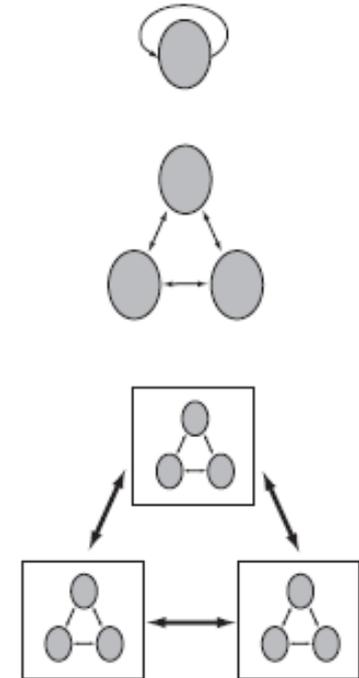
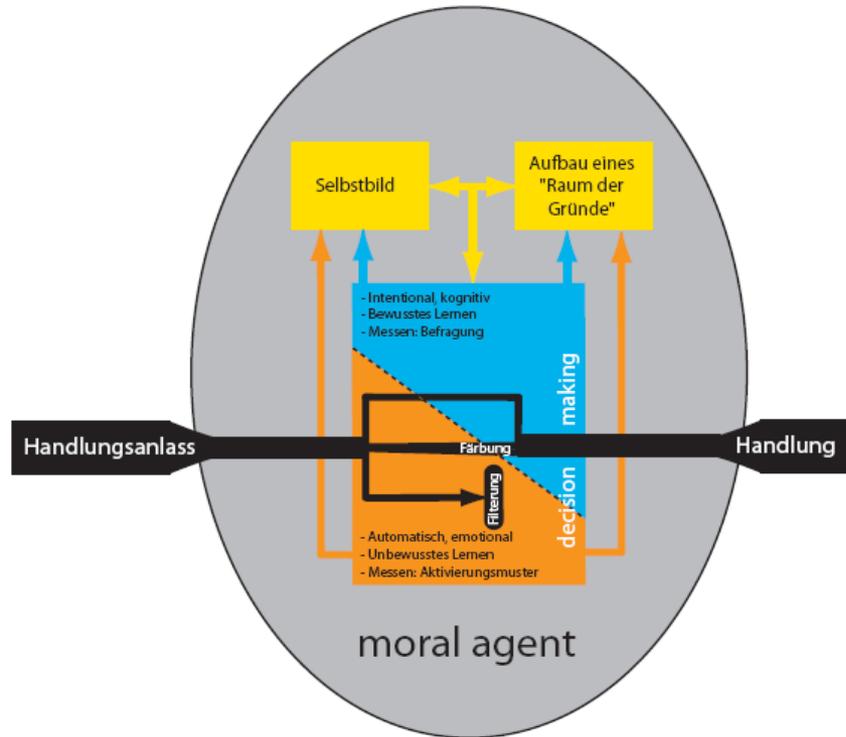
Generell wird *agency* als die Fähigkeit einer Entität verstanden, ihre Umwelt wahrzunehmen, aufgrund dieser Wahrnehmungen und inneren Zuständen Entscheidungen zu fällen und basierend auf diesen Entscheidungen auf die Umwelt handelnd einzuwirken, ohne dass diese Prozesse einer strikten Kontrolle von Aussen unterliegen.

## Merkmale eines **Agent**:

- Autonomie
- Urheberschaft
- Wertsetzung
- Zielsetzung aufgrund Wertsetzung
- Zielorientierung im Handeln
- Interaktivität mit anderen *agents*
- Reaktivität (kurze Zeitskala)
- Adaption (lange Zeitskala)



# Der moral agent / Ebenen der Interaktion



Die Struktur von *moral agency* lässt sich in vier Komponenten gliedern:

**Wahrnehmung:** Ein für den betreffenden *moral agent* sinnlich erfahrbares raumzeitliches Ereignis

**Entscheidung:** Ein intentionaler und ein automatisch ablaufender Prozess eines *decision making*.

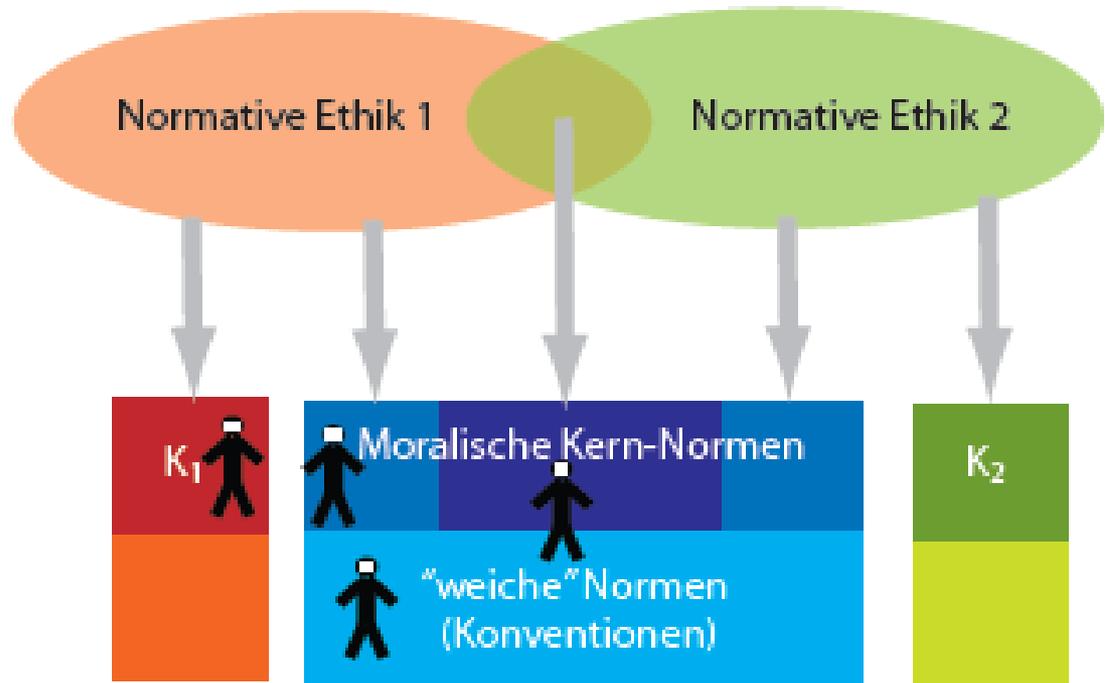
**Handlung:** Das Wirken des *agent* in der Raumzeit.

**Begründung:** Der Komplex von Begründungen, welche der betreffende *moral agent* für die Rechtfertigung seiner Handlung angibt (angeben würde).

# Moral und Ethik

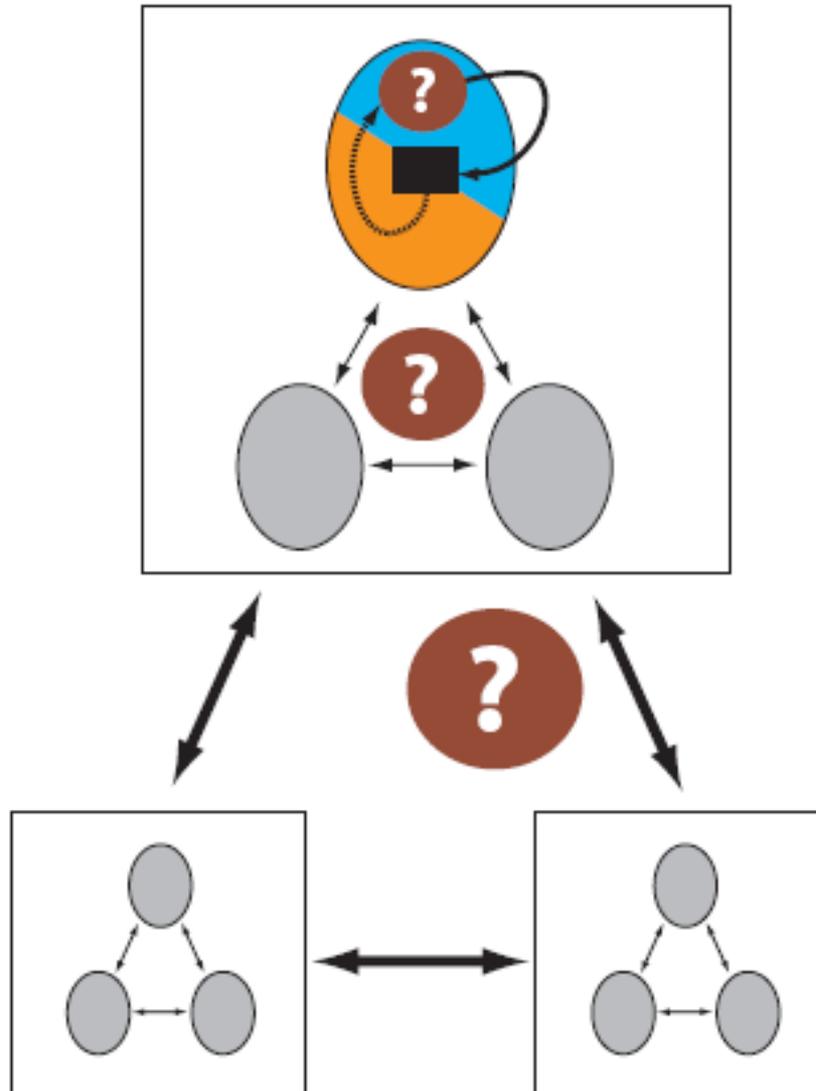
Empirische Zugangsweisen zu *moral agency* bedürfen einer gewissen inhaltlichen Festlegung. Doch es gibt verschiedene Varianten von Festlegungen:

Normative  
Ethik



Deskriptive  
Ethik (Moral)

# Wo wirkt „technisierte“ Moral Agency



# Beeinträchtigung von Agency

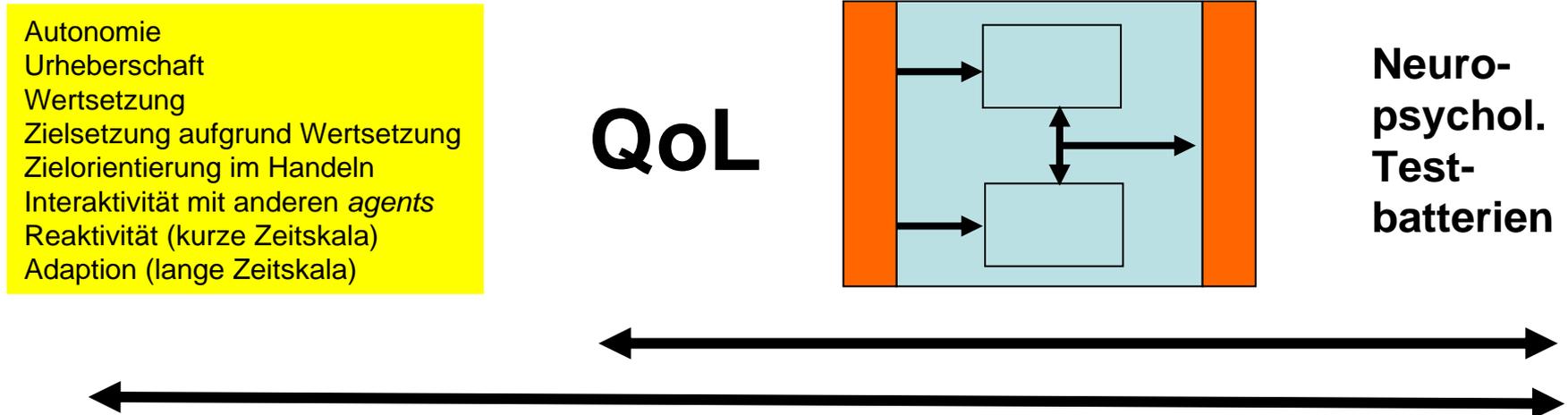
		Kontextbedingung		
		Individuum	Beziehung	Gesellschaft
Struktur von moral agency	<b>Wahrnehmung</b>	Unfähigkeit, sich emotional von Wahrnehmungen berühren zu lassen. Fehlgeleitete emotionale Belegung von Wahrnehmungen. Probleme, Kausalitäten eines Wahrnehmungskomplexes zu erkennen.  <b>Angststörungen, Autismus</b>	Beziehungs-spezifische Wahrnehmungsstörungen, eingeschliffene Wahrnehmungsdefizite in gewissen Beziehungskontexten.	Frage des Zugangs zur relevanten Information, kollektive Ausblendung von Missständen, die Anlass zu moralischen Debatten liefern könnten.
	<b>Entscheidung</b>	Unvermögen zur Wertsetzung, Priorisierung von Werten.  <b>Depression</b>	Einschränkung der Optionen durch beziehungsinterne Tabus.	Fragen der Art der Entscheidungsverfahren und der Verteilung von Macht und Kontrolle innerhalb dieser Verfahren.
	<b>Handlung</b>	Behinderungen, welche das Einwirken auf die Welt generell erschweren (Bewegungs-/Kommunikationsstörung).  <b>Tourette-Syndrom / Zwangsstörungen</b>	Peer pressure.	Fragen des empowerment, Kontrolle über Umsetzung, Reichweite dieser Kontrolle.
	<b>Begründung</b>	Kognitive Störungen (Gedächtnis, Zugang zur eigenen Biografie).  <b>(Autismus? / Depression?)</b>	Abgekapselte ideologische Systeme (z.B. Sekten).	Problem von stabilen kollektiven Irrationalitäten.

# Agency als Thema der Erfolgskontrolle?

Untersuchung von 34 Tests, die bei der Beurteilung des Erfolgs von DBS bei psychiatrischen Erkrankungen eingesetzt wurden, hinsichtlich ihres Gehalts an impliziten, *agency-relevanten* Aspekten untersucht

Bezug einzelner Aspekte: 11 (7 QoL, 2 Depr.)  
Gewisser Bezug: 10  
Kaum Bezug: 13

*Interessant ist die mehrfach aufgetretene Beobachtung, dass Patienten offenbar nicht bemerken konnten, ob der Stimulator aktiv ist oder nicht – obgleich sie Veränderungen ihres Zustandes bemerken.*



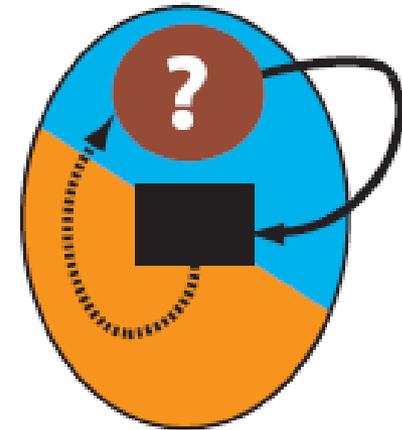
# Problemfeld 1: Agency testen (1)



- Zwei „Testobjekte“: - Transformationsprozess  
- Fehlende Übereinstimmung

# Problemfeld 1: Agency testen (2)

**Eine Möglichkeit:** Erleben von *agency* durch den Patienten unter Kontrolle/Beobachtung innerer Parameter (DBS), äusserer Parameter (Wirkorte von *agency*) und einem operationalisierbaren Begriff von *agency* untersuchen.



Komponenten	Ansatzpunkte von moral agency	Konkrete Tests im akuten und chronischen Szenario.
<b>Wahrnehmung</b>	Bewertung des „moralischen Gehalts“ bestimmter (affektiver) Bildstimuli, akustischer Stimuli.	Einsatz von kürzlich entwickelten Empathie-Tests.
<b>Entscheidung</b>	Einsatz der bekannten moralischen Dilemmata der Moralpsychologie	Trolley-Dilemma, weitere?
<b>Handlung</b>	Bewertung von sich selbst, der moralische oder unmoralische Handlungen vollzieht.	Erfragen von biografischen Erinnerungen von Scham/Schuld?
<b>Begründung</b>	Ermitteln des eigenen Wertprofils und dessen Kohärenz	Anhand ausgewählter Entscheidungssituationen Skizzen anfertigen, welche Gewichtungen von Gründen zeigen.

# Problemfeld 2: Agency verändern?

Könnte DBS als Instrument genutzt werden, um „moralische Pathologien“ zu beheben – also Personen zu behandeln, welche Defizite bei der Wahrnehmung moralrelevanter Fähigkeiten (z.B. Empathie) oder der moralischen Kognition (Entscheidungsfindung in sozialen Kontexten) zeigen und moralisch unangemessen handeln (z.B. unmotivierter Gewaltausbrüche)?

Derzeit wird eine derart negative Beeinflussung von *agency* ja primär als unbeabsichtigte Nebenfolge von DBS angesehen (Studien von Bothe, Münster!).

NoE liefert derzeit kaum Ansatzpunkte für Bestimmung von Stimulationsparametern. Weiterentwicklung dieser Theorie hat zahlreiche methodischen und normativen Schwierigkeiten - ist aber durchaus auch von starken Motiven begleitet:

**Hier zeichnet sich ein Problem ab!**  
**(Zielvorgabe der zu erfüllenden Kern-Normen).**