Will Artificial Emotional Agents Show Altruistic Punishment In The Public Goods Game?

Dirk M. Reichardt

BA Stuttgart
University of Cooperative Education
D-70180 Stuttgart, Germany
reichardt@ba-stuttgart.de

BERUFSAKADEMIE STUTTGART
University of Cooperative Education

This talk is about ...

modeling an emotional agent in the PGG

Will

Artificial Emotional Agents

Show

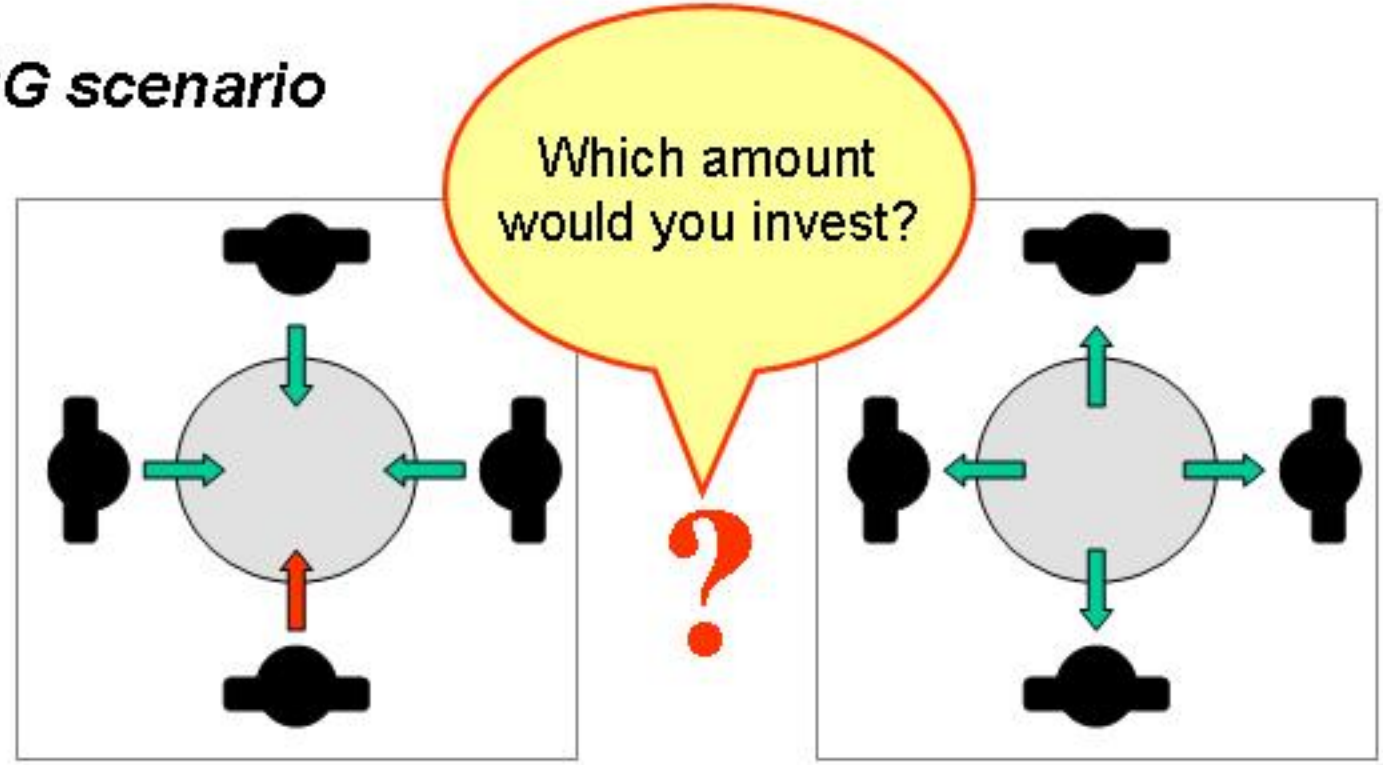
Altruistic Punishment

In The Public Goods Game?

What is the role of altruism and what's the link to emotion?

game theory and rational behavior ... the *scenario* for the emotional agent

The PGG scenario



- 4 participants
- each participant gets 20 Euros
- choice to invest into a public project (0-20 Euros)
- the project is bearing 60 % interest
- the invested amount plus interest is equally split

The PGG scenario ... simplified

How do we get both players in a situation in which both are better off?

NASH-EQUILIBRUM

Dominant Strategy:
"don't pay!"

Agent A \ Agent B	0	20
0	20 / 20	36 / 16
20	16 / 36	32 / 32

A **rational agent** would not invest anything in a public project !

Poor society!

BERUFSAKADEMIE
STUTTGART
University of Cooperative Education

[the public goods game]

The PGG scenario ... *with punishment*

Less than before?
So would you punish the „free rider“?

- invest 0-10 Euros to punish
- punished player must pay ***3 times**** the invested sum

	Player A	*Player B*	*Player C*	*Player D*
investment	*0*	*20*	*20*	*20*
outcome	*44*	*24*	*24*	*24*
punishment	*-90*	*0*	*0*	*0*
punishment cost	*0*	*-10*	*-10*	*-10*
outcome	*-46*	*14*	*14*	*14*

** scalable in other experiments*

The iterated PGG scenario ... with punishment

- iteration of the situation makes punishment *rational*

- ***why***? The free rider *learns* that the strategy is wrong

Fehr and Gächter setup:

- exchanging the group participants after each iteration

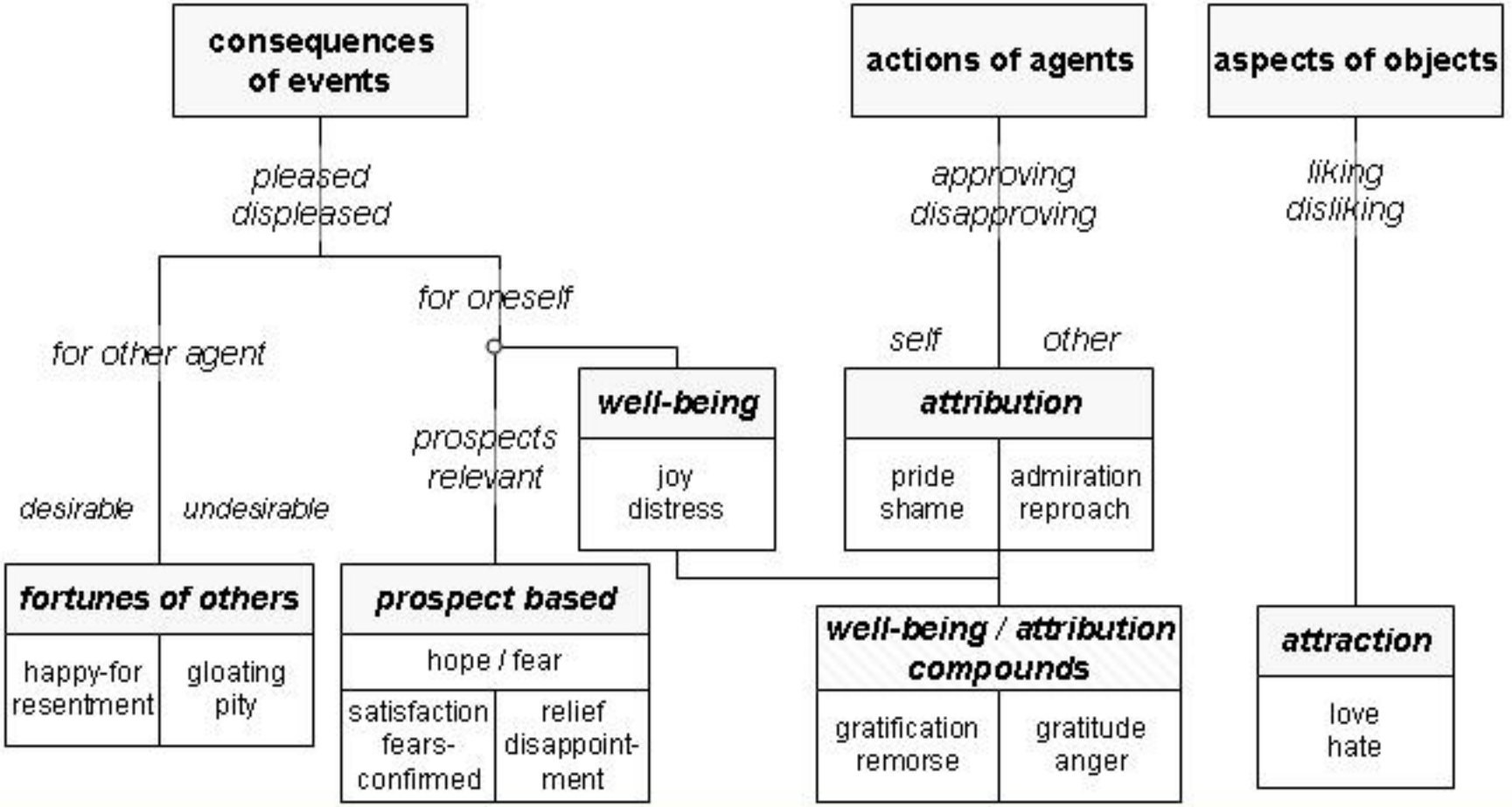
→ If the free rider learns from punishment,
the punisher does *not* get a return!

but ... others will benefit from the change of behavior!

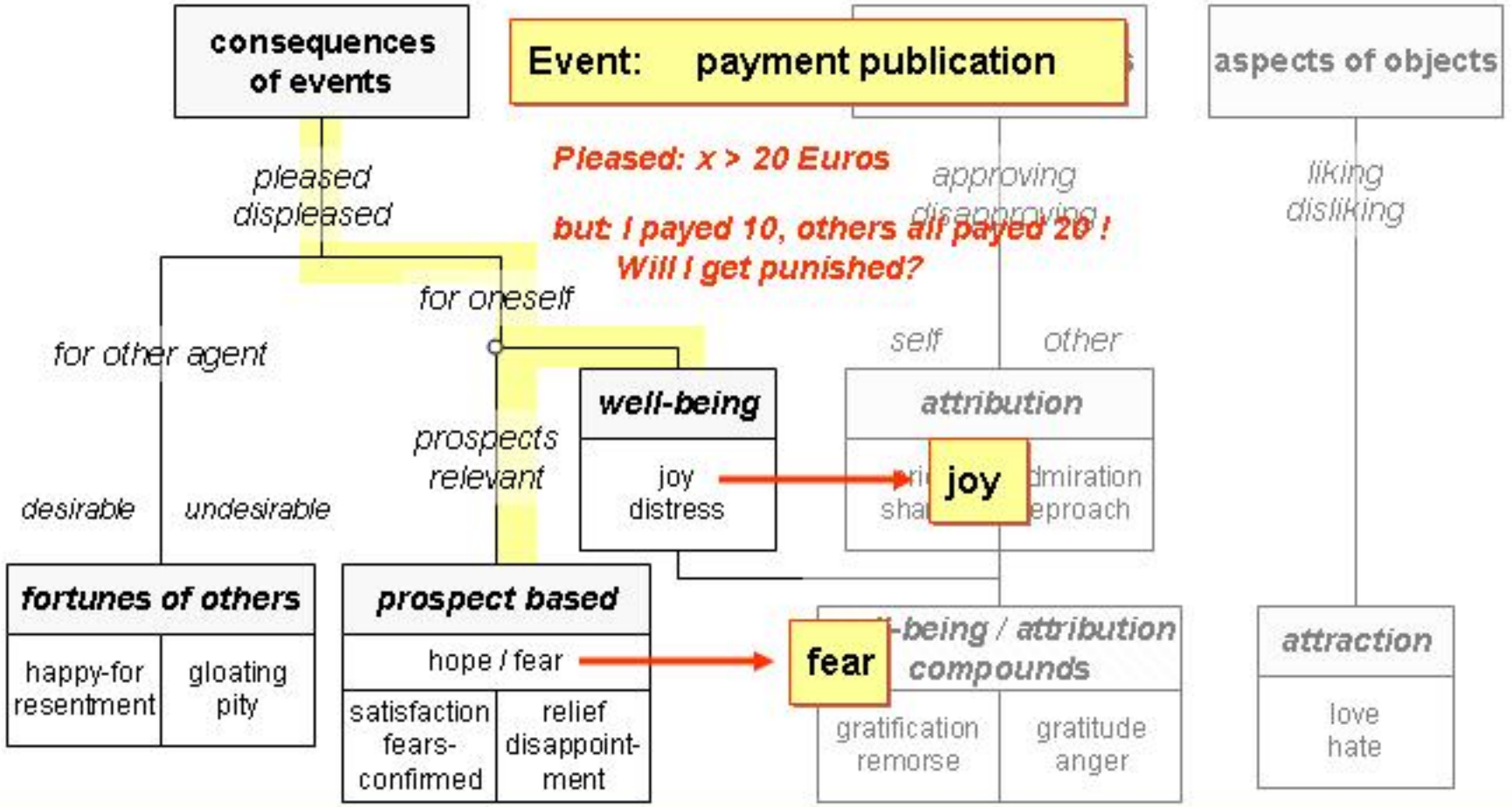
↓

Altruism

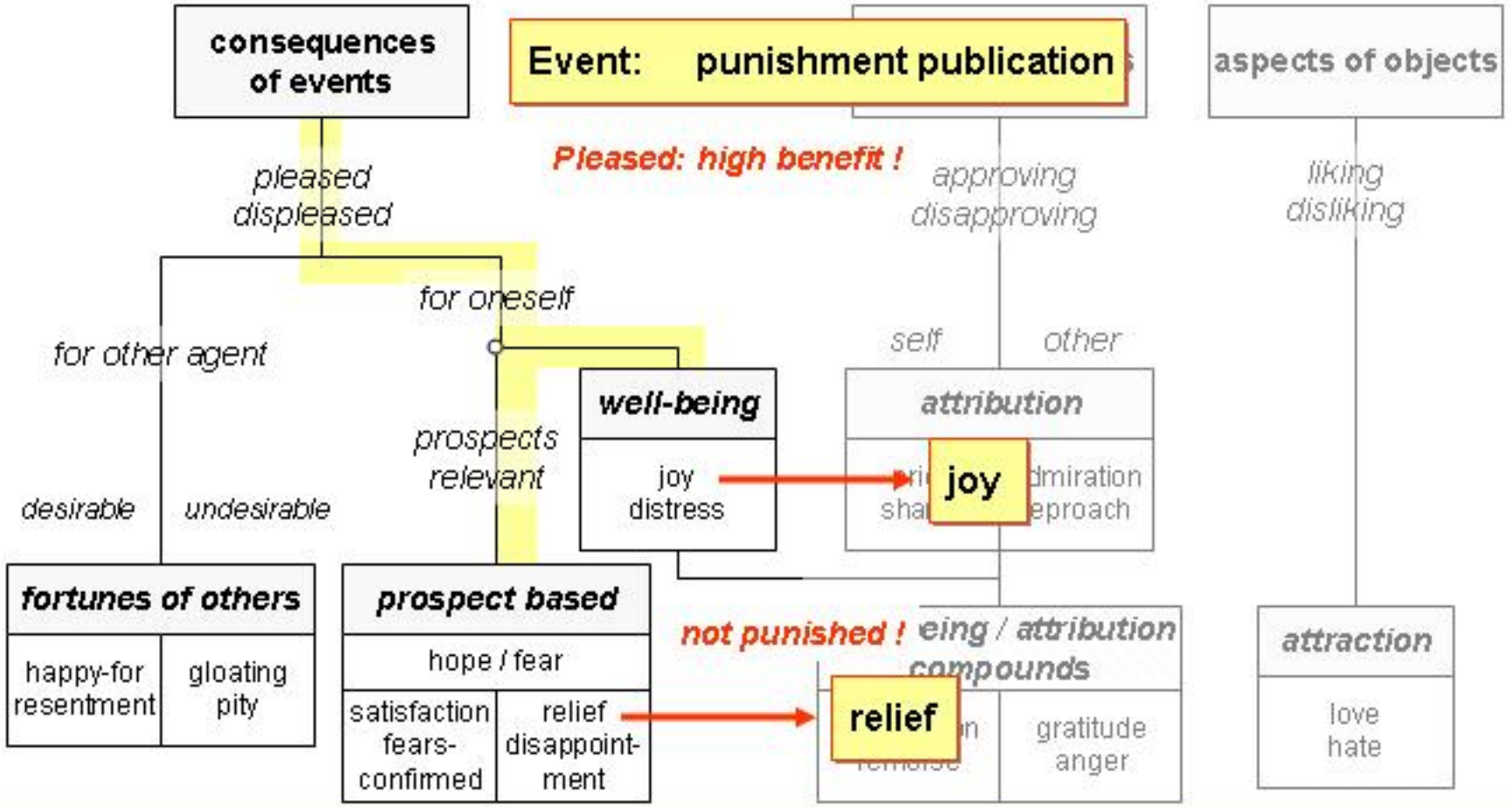
The OCC model



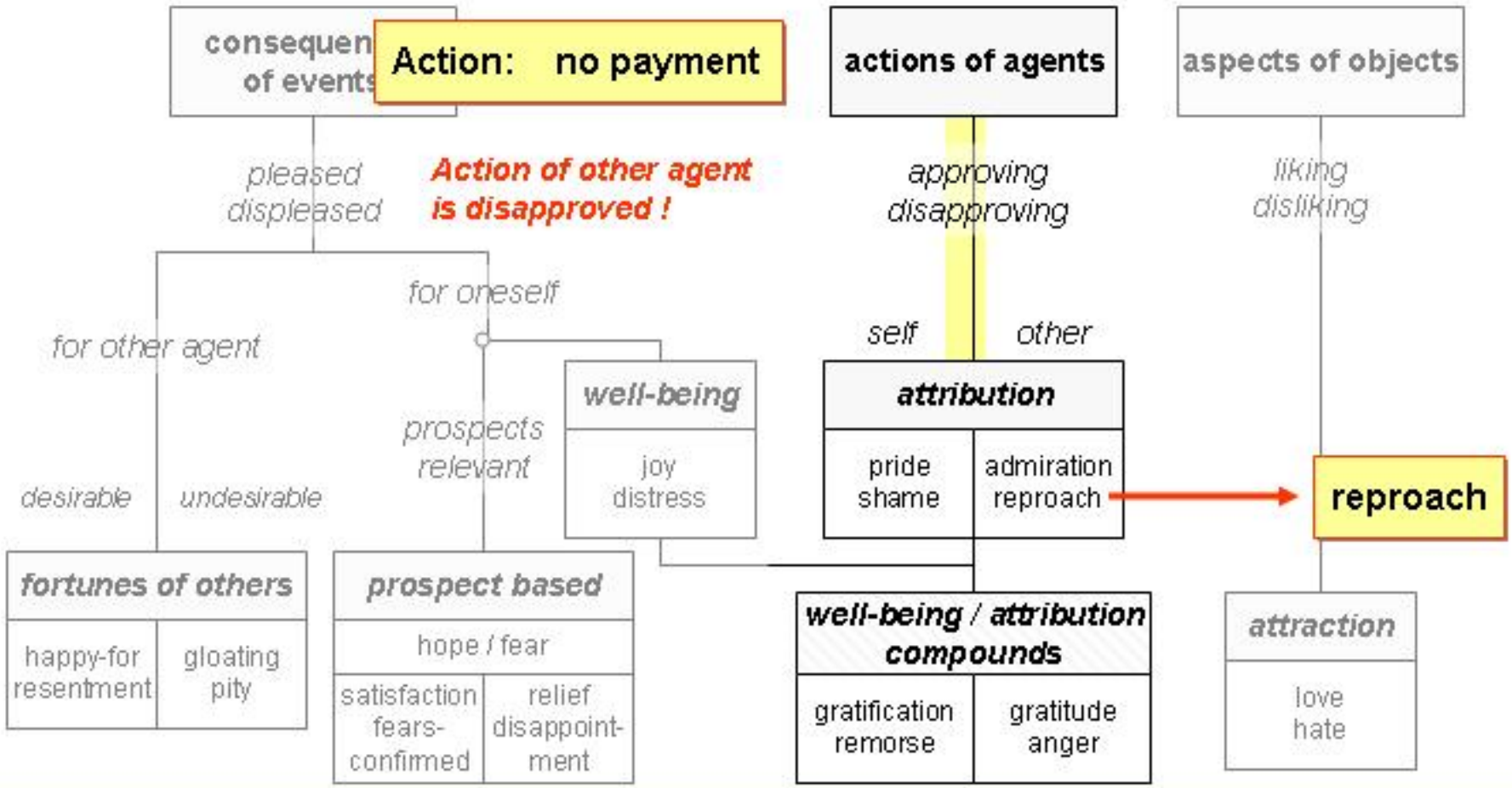
The OCC model



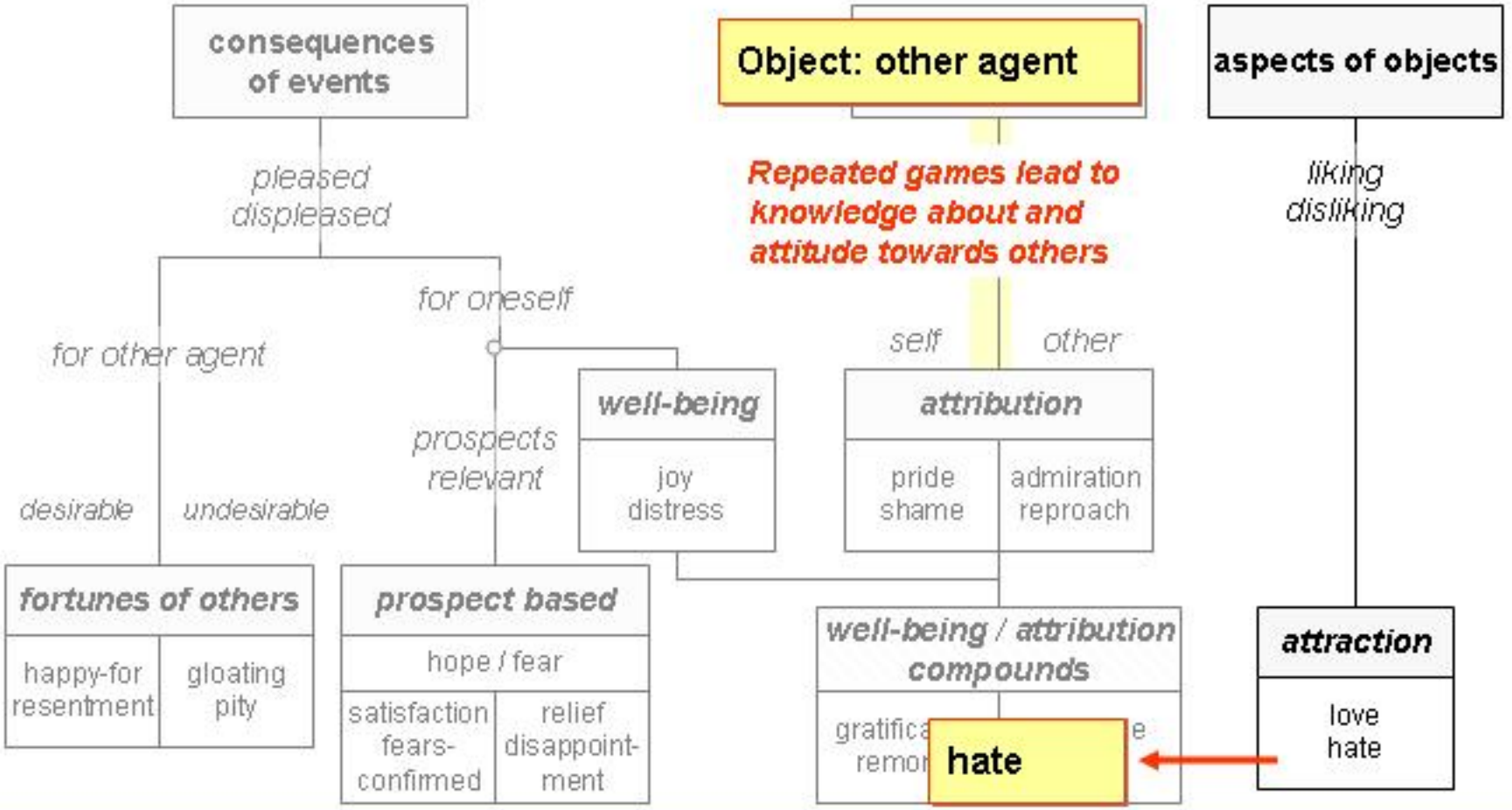
The OCC model



The OCC model



The OCC model

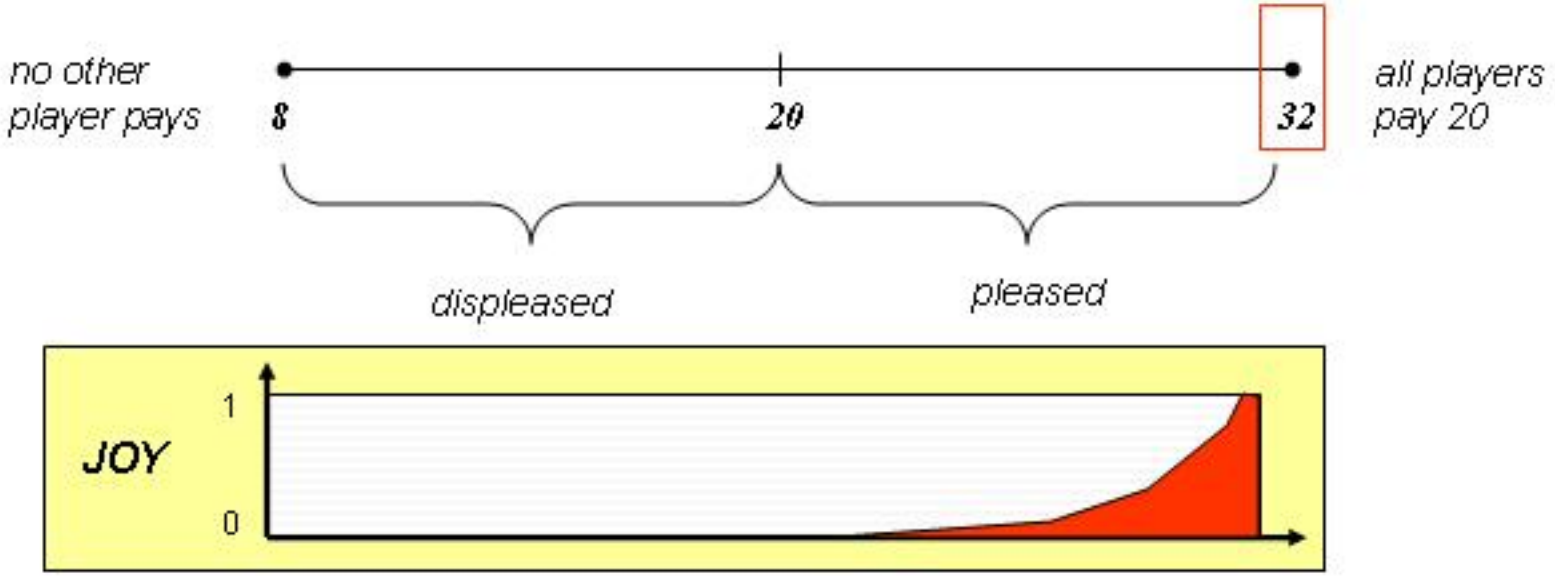


Emotion intensities

Let's play !

Goals: a) get as much money as possible
b) get more money than the others
c) ...

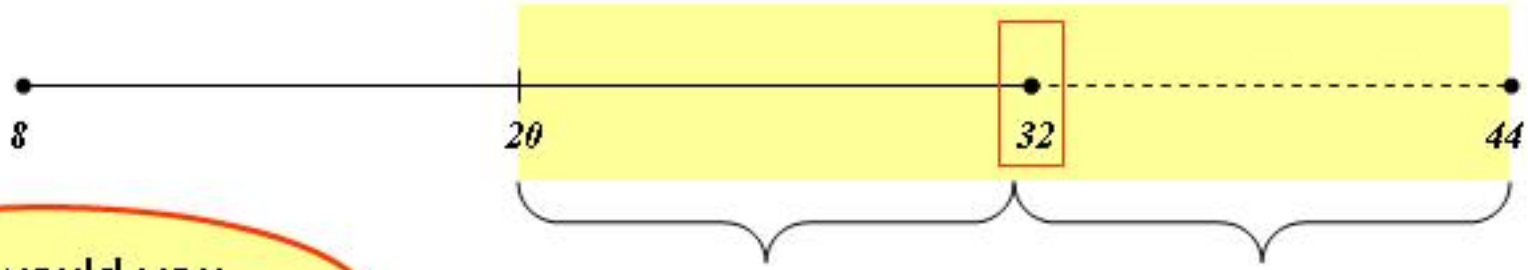
Event: **32 Euros as result !**



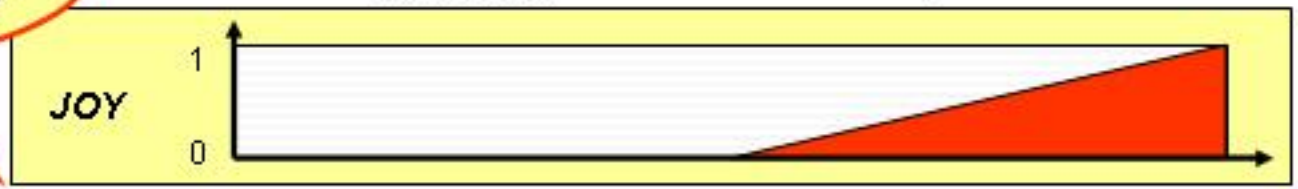
Emotion intensities

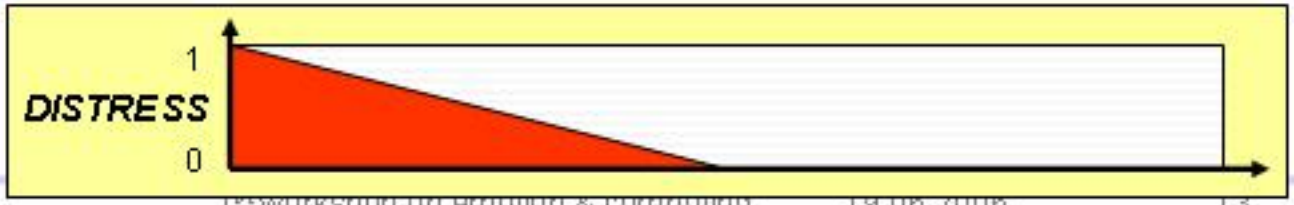
Now: assume the player invested nothing ...

Event: **32 Euros as result !**



How would you feel about the event?





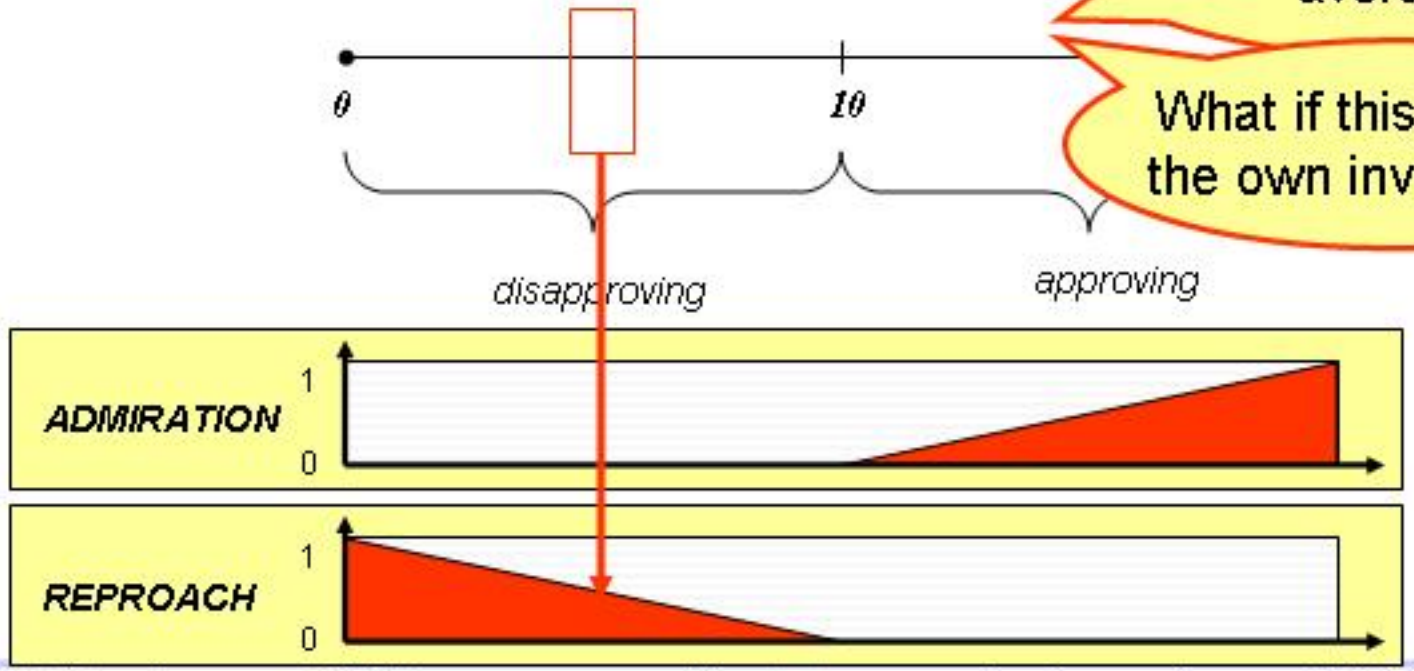
Emotion intensities

Standards a) behavior of others/own - what is typical
b) individual moral standard
c) ...

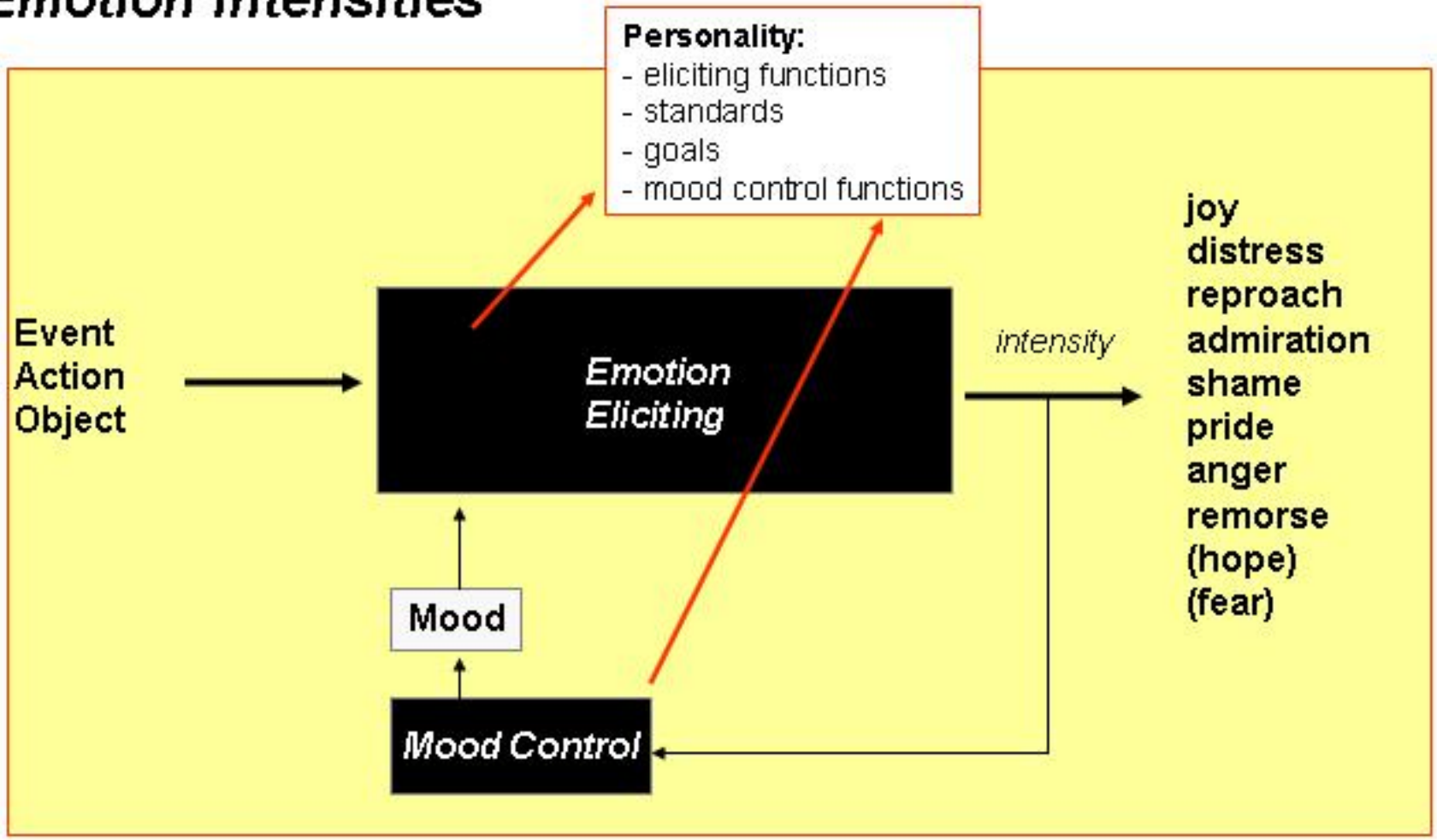
Action: **Investment of 5 Euros by player 2!**

What if this is above average?

What if this is above the own investment?



Emotion intensities



Personality

Eysenck introvert vs. extrovert / stable vs. non-stable

here: openness: rigidly sticking with a moral standard or strategy vs. exeptions

energy: fast vs. slow mood changes, tendency to emotion based reactions

↓

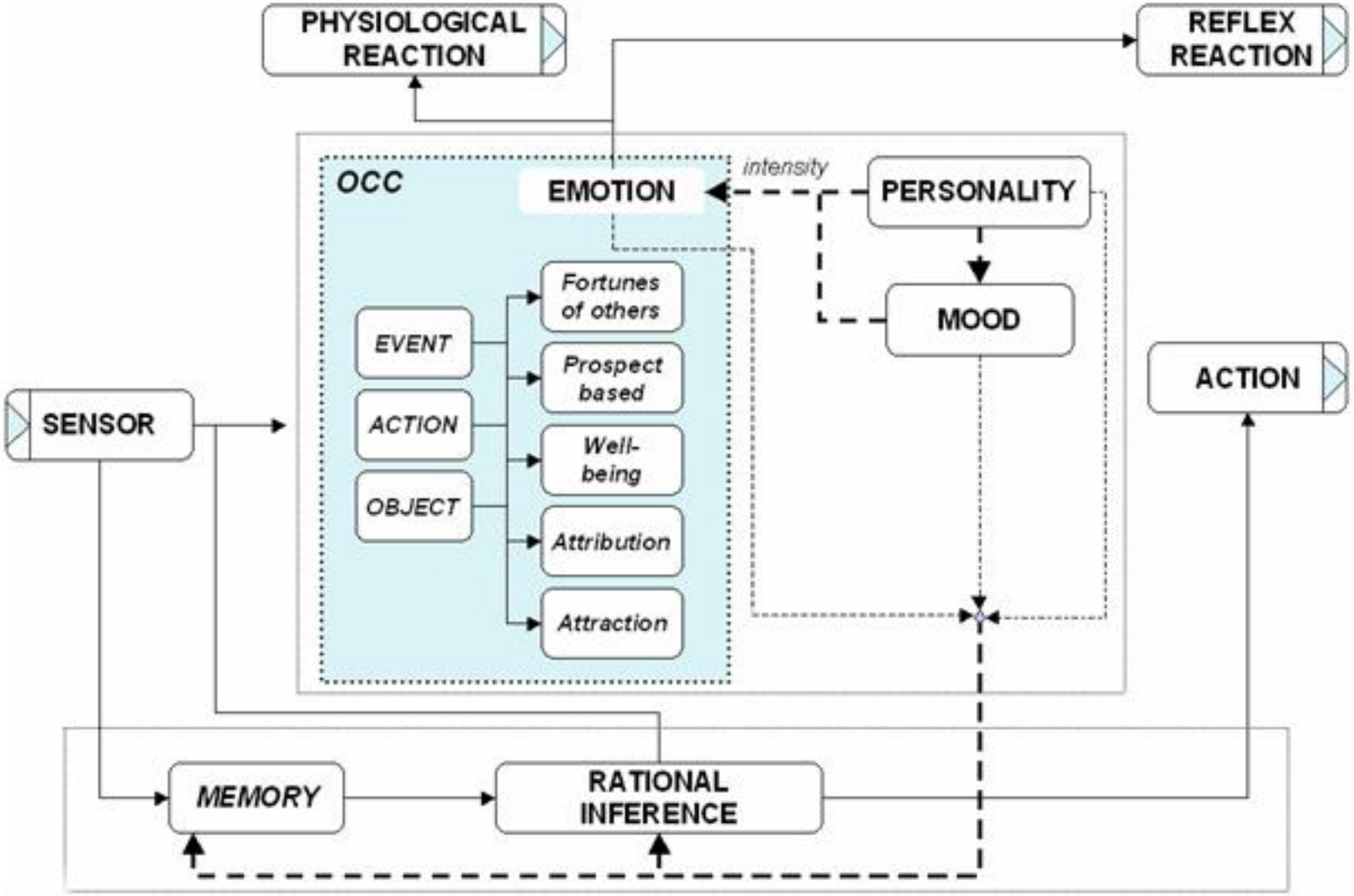
parameters for eliciting functions

function selection depends on *goal* as well

Hardcoded in this „toy example“

BERUFSAKADEMIE STUTTGART
University of Cooperative Education

[architecture]



Decision making

- integrating emotion in a ruleset ***as facts*** (observations)
- alternative: adding a goal of ***well being***

investment

<last_invest = x> AND <pride> AND <joy> THEN <action = x>

<last_invest = x> AND <shame> AND <distress> THEN <action = x++>

<last_invest = x> AND <pride> AND <distress> THEN <action = x- ->

<last_invest = x> AND <shame> AND <joy> AND <likes y> THEN <action = x++>

punishment

<pride> AND <distress> AND <reproach y> THEN <punish y 1>

Conclusion

Now, will
Artificial Emotional Agents
Show
Altruistic Punishment
In The Public Goods Game,
or not?

- well, they can be told to do so ...
- better question: will they still do so, once they mimic rational and emotional behavior?
- which form of agent personality will survive the pgg society?
- model understanding: which parts of OCC are hard to apply to the situation?

Next steps:

- integrate more features in the model
- complete the implementation
- add facial expression module
- experiments with human pgg groups
- experiments with agent populations